

University of Groningen

Attitude en handelen van basisschoolleerkrachten met betrekking tot het verbeteren en borgen van leerlingresultaten

Deunk, M.; Doolaard, S.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2013

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Deunk, M., & Doolaard, S. (2013). *Attitude en handelen van basisschoolleerkrachten met betrekking tot het verbeteren en borgen van leerlingresultaten*. s.n.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Attitude en handelen van basisschoolleerkrachten met betrekking tot het verbeteren en borgen van leerlingresultaten

Resultaten van Vragenlijst Streef 2010 en 2012

Dr. Marjolein Deunk
Dr. Simone Doolaard

Attitude en handelen van basisschoolleerkrachten met betrekking tot het verbeteren en borgen van leerlingresultaten

Dr. Marjolein Deunk · Dr. Simone Doolaard

Attitude en handelen van basisschoolleerkrachten met betrekking tot het verbeteren en borgen van leerlingresultaten

Resultaten van Vragenlijst Streef 2010 en 2012

Rapportage BOPO-programmalijn Onderwijskwaliteit PO
Deelproject 2 'Nadruk op basisvaardigheden' (413-09-062)

Dr. Marjolein Deunk & Dr. Simone Doolaard
Groningen: GION/RUG
Februari 2013

ISBN 978-90-6690-578-8

2013. GION, Gronings Instituut voor Onderzoek van het Onderwijs

Niets uit deze uitgave mag worden verveelvoudigd en/of openbaar gemaakt door middel van druk, fotokopie, microfilm of op welke andere wijze dan ook zonder voorafgaande schriftelijke toestemming van de directeur van het instituut.

No part of this book may be reproduced in any form, by print, foto print, microfilm or any other means without written permission of the director of the instituut.

Inhoud

Managementsamenvatting	1
1 Inleiding	5
2 Methode	9
2.1 Respondenten	9
2.2 Analyseplan	11
3 Resultaten	13
3.1 Opbrengstgericht werken	13
3.1.1 Eindniveaus en prestaties	13
3.1.2 Referentieniveaus	16
3.1.3 Toetsgebruik en rapportage	20
3.2 Differentiatie	25
3.3 Professionaliseren en samenwerken in het team	26
3.4 Attitude en handelen van schoolleiders en intern begeleiders	28
3.4.1 Tevredenheid prestaties en explicatie van doelen	28
3.4.2 Een opbrengstgerichte attitude	30
3.4.3 Bekendheid referentieniveaus	30
3.5 Opbrengstgerichtheid van teams en leerwinst bij leerlingen	34
3.5.1 Clusters leerkrachtenteams	34
3.5.2 Resultaten per jaargroep	38
4 Conclusie en discussie	41
Literatuur	47
Bijlage	48

Managementsamenvatting

Binnen onderzoeksproject Streef is in de schooljaren 2010-2011 en 2011-2012 een aantal schoolverbeteringstrajecten uitgevoerd, waarin opbrengstgericht werken en het stellen van doelen centraal staan. Met een leerkrachtvragenlijst is de attitude en het handelen rondom deze onderwijspraktijken en het effect van de deelname aan Streef hierop in kaart gebracht. De attitude en het handelen met betrekking tot prestatiedoelen, referentieniveaus, toetsgebruik, differentiatie en professionalisering van leerkrachten en van schoolleiders en intern begeleiders komen aan de orde. De nadruk in de rapportage ligt hierbij op de 345 leerkrachten die de vragenlijst zowel in 2010 als in 2012 hebben ingevuld, waarvan ongeveer de helft (47%) heeft deelgenomen aan een van de Streefprojecten. Om het effect van de schoolverbetertrajecten in kaart te kunnen brengen worden de leerkrachten die hebben deelgenomen aan Streef vergeleken met de leerkrachten die niet aan een van de deelprojecten hebben meegenomen. Daarnaast worden de attitudes en werkwijzen van Streef-leerkrachten vergeleken met die van schoolleiders en directeurs in 2012, om in kaart te kunnen brengen in hoeverre schoolleiders en directeurs een voortrekkende rol kunnen spelen in schoolverbetering. Tenslotte is het effect van de opbrengstgerichtheid van leerkrachtenteams op de leerwinst van leerlingen onderzocht.

Samengevat kan gesteld worden dat de schoolverbetertrajecten van Streef hebben geleid tot meer en intensiever gebruik van toetsgegevens en -rapportages. Daarnaast hebben de trajecten de attitudes van leerkrachten beïnvloed, vooral op het gebied van opbrengsten en professionalisering. Beter toetsgebruik en een positievere houding tegenover aspecten van opbrengstgericht werken vormen een belangrijke basis voor kwaliteitsverbetering in het onderwijs en zijn hiermee een eerste stap in de richting van prestatieverhoging van leerlingen. Daarmee zijn scholen er echter nog niet: voor het vergroten van de leerwinst van leerlingen lijken een voortrekkende rol van de schoolleider en intern begeleider en vooral een opbrengstgericht, samenwerkend en differentiërend team essentieel. De meest opvallende resultaten worden hieronder toegelicht.

Leerdoelen en opbrengsten

Bij vragen naar het vaststellen van leerdoelen vallen met name de problemen met begrijpend lezen op. Leerkrachten geven aan het minst tevreden te zijn over het begrijpend leesniveau van hun leerlingen. Bovendien stellen zij relatief vaak geen eindniveau vast dat zij willen behalen aan het eind van het schooljaar. Dit is een gemiste kans, aangezien expliciete leerdoelen kunnen bijdragen aan prestatieverbetering. Slechts ongeveer de helft van de leerkrachten geeft aan de leerdoelen voor de verschillende vakken samen met het team te expliciteren. De schoolleiders en IB-ers bevestigen dit beeld, al rapporteren zij iets meer teamoverleg. De Streef-interventies dragen iets bij op dit gebied: leerkrachten stellen na

deelname aan Streef vaker rekendoelen vast in overleg met de intern begeleider of het team dan voor de interventie, maar verdere verbetering is wenselijk.

De Streef-interventies hangen verder samen met de opbrengstgerichte attitude van leerkrachten. Leerkrachten die deel hebben genomen aan een van de deelprojecten zijn in de periode 2012-2012 significant positiever geworden over de haalbaarheid en wenselijkheid van het verbeteren van leerlingprestaties en hebben daardoor in 2012 een significant positievere attitude over opbrengsten dan de leerkrachten in de controlegroep. Schoolleiders en intern begeleiders zijn daarbij in 2012 nog positiever over opbrengsten en basisvaardigheden dan de leerkrachten.

Referentieniveaus

De referentieniveaus zijn een door de overheid ingevoerd middel ter stimulering van een opbrengstgerichte werkwijze op school. In de periode van 2010 tot 2012 is de bekendheid met referentieniveaus bij alle leerkrachten significant toegenomen, vooral bij de leerkrachten die hebben deelgenomen aan een van de Streef-interventies. Bovendien zijn alle schoolleiders en intern begeleiders in 2012 bekend met de referentieniveaus. Dit wil echter niet zeggen dat de referentieniveaus ook leven: nog geen derde van de leerkrachten praat erover op school. Bijna niemand van de leerkrachten, schoolleiders en intern begeleiders verwacht een negatief effect van de referentieniveaus, maar ruim de helft verwacht geen effect of weet het niet. Bovendien zijn de positieve verwachtingen die leerkrachten hadden in 2010 getemperd in 2012, wellicht door de grotere bekendheid met de referentieniveaus. Desondanks staan de leerkrachten redelijk positief tegenover de referentieniveaus. Zij zien de meerwaarde vooral in het expliciteren van doelen en het gebruik bij differentiatie.

Toetsrapportages

Leerkrachten, vooral degenen die hebben deelgenomen aan een Streef interventie, maken in 2012 gebruik van meer toetsrapportages en voeren uitgebreidere analyses uit op de toetsgegevens dan in 2010. De Streef-leerkrachten maken in 2012 bovendien significant meer gebruik van uitgebreide analysemogelijkheden dan de leerkrachten in de controlegroep. Ook de schoolleiders en intern begeleiders maken ruimschoots gebruik van de analysemogelijkheden van de Cito LOVS toetsen. De leerkrachten in de Streef-groep zijn daarnaast in de periode 2010-2012 significant meer gaan reflecteren op hun onderwijspraktijk en de eerdere prestaties van de leerlingen naar aanleiding van toetsresultaten. De leerkrachten in de Streef-projecten hebben zich dus positief ontwikkeld in gebruik, analyse en interpretatie van toetsgegevens en zijn hiermee, in ieder geval wat betreft toetsgebruik, meer opbrengstgericht gaan werken.

Handelen: differentiatie en samenwerking

Differentiatie is een essentieel onderdeel van opbrengstgericht werken. Er zijn nog verbeteringen nodig op gebied van het optimaal gebruik maken van verschillende prestatieniveaus, het aanbieden van verschillende verwerkingsactiviteiten na een toets en de houding tegenover de wenselijkheid en haalbaarheid van differentiatie. Alle leerkrachten scoren in 2010 en 2012 laag tot gemiddeld op de verschillende differentiatieonderdelen, ook degenen die deel hebben genomen aan een van de Streef projecten. Zij zouden hier dus meer ondersteuning bij kunnen gebruiken. Ook zouden zij meer kunnen samenwerken. De meeste samenwerking betreft het uitwisselen van lesmateriaal of het praten over de ontwikkeling van nieuw lesmateriaal met collega's. Lesobservaties zijn een belangrijke bron van informatie bij professionele ontwikkeling, maar slechts een-derde van de leerkrachten observeert wel eens bij een collega in de les.

Opbrengstgericht werken binnen leerkrachtenteams

De leerkrachten die de vragenlijst hebben beantwoord, vormen teams in hun scholen. Er zijn *opbrengstgerichte* teams, die veel samenwerken, reflecteren en differentiëren en die veel waarde hechten aan rekenen en taal. *Individualistische* teams hechten wel veel waarde aan de basisvaardigheden en voeren uitgebreide rapportages uit, maar werken weinig samen en reflecteren en differentiëren weinig. *Behoudende* teams hechten weinig waarde aan innovatie en professionalisering, zijn gemiddeld in mate van samenwerking en aandacht voor prestaties en differentiëren veel. *Niet opbrengstgerichte* teams tenslotte hechten weinig belang aan basisvaardigheden en opbrengsten. De attitudes en het handelen van de leerkrachten in de schoolteams zijn van invloed op de leerwinst van de leerlingen. Leerlingen van *opbrengstgerichte* en *behoudende* leerkrachtenteams boeken meer vooruitgang dan leerlingen van *individualistische* en *niet opbrengstgerichte* teams. De *opbrengstgerichte* en *behoudende* teams hebben gemeen dat zij relatief vaak differentiëren in de klas. Behoudende teams staan echter minder open voor onderwijsvernieuwingen en zien minder noodzaak zichzelf te blijven ontwikkelen op professioneel vlak. Het valt op dat het maken van uitgebreide toetsrapportages, zoals de *individualistische* teams doen, niet genoeg is om extra leerwinst te boeken. Leerkrachten moeten ook naar de informatie uit de rapportages *handelen* en moeten met elkaar als team werken aan en reflecteren op hun onderwijs. Dit geeft het belang aan van ondersteuning bij het schoolbreed implementeren van opbrengstgericht werken, een van de uitgangspunten van de Streefprojecten.

1 Inleiding

Om prestaties van zwakke, maar ook van sterke basisschoolleerlingen, op gebied van taal en rekenen te verbeteren, wordt de laatste jaren veel aandacht besteed aan opbrengstgericht werken. Het vaststellen van nationale leerstandaarden in de vorm van referentieniveaus is hier onderdeel van (Expertgroep Doorlopende Leerlijnen Taal en Rekenen, 2009). Bij het opbrengstgericht werken worden verwachte en behaalde prestaties van leerlingen gebruikt in een cyclus van plannen, uitvoeren, evalueren en bijstellen van onderwijs (Inspectie van het Onderwijs, 2010; Ledoux, Blok, & Boogaard, 2009; Schildkamp & Kuiper, 2010; Visscher & Coe, 2003). Deze onderwijspraktijk bevat elementen die een positief effect hebben op de leerprestaties, zoals het regelmatig evalueren van prestaties en het hebben van hoge verwachtingen van leerlingen (Scheerens & Bosker, 1997).

Het expliciteren van doelen of verwachtingen is een belangrijk onderdeel van opbrengstgericht werken. Doelen en verwachtingen zijn aan elkaar gerelateerd: als een verwachting wordt uitgesproken en onderbouwd, en de leerkracht streeft ernaar het verwachte niveau te behalen, dan functioneert de verwachting als doel. Het doel zou daarbij wel hoger moeten liggen dan het niveau dat verwacht wordt op basis van de ‘normale’ groei van de leerling. Er zijn verschillende onderzoeken gedaan naar het effect van leerkrachtverwachtingen. Het bekendste is waarschijnlijk *Pygmalion in the classroom* (Rosenthal & Jacobson, 1968), waarin de onderzoekers laten zien dat de verwachtingen van de leerkracht zijn of haar gedrag bepalen, waardoor de kans groot is dat deze verwachtingen ook daadwerkelijk uitkomen. Uit latere onderzoeken naar het effect van verwachtingen blijkt dat uiterlijk, gedrag en sociale achtergrond van leerlingen meespeelt bij de prestatieverwachtingen die leerkrachten van hen hebben, maar dat dit afneemt als er meer objectieve prestatiegegevens, zoals toetsscores, voorhanden zijn (Hattie, 2009). Andere onderzoekers hebben bestudeerd hoe leerkrachten hun gedrag aanpassen aan de verwachting die zij van de leerlingen hebben. Hieruit blijkt dat de verwachting beïnvloedt welke vragen de leerkracht aan de leerling stelt, hoe vaak en wanneer een leerling de beurt krijgt, wat voor soort feedback een leerling krijgt en welke lesstof de leerling krijgt aangeboden (Jungbluth, 2003; 2012; Weinstein, 2002). Daarnaast zijn er aanwijzingen dat de verwachtingen die leerkrachten (en ouders) van leerlingen hebben beïnvloeden wat leerlingen van zichzelf verwachten (Weinstein, 2002). Het is kortom ongunstig voor leerlingen als hun leerkracht een relatief lage verwachting van hen heeft en gunstig als die verwachting relatief hoog is.

Opbrengstgericht werken is dus meer dan het uitvoeren van analyses op afgenomen toetsen alleen (de ‘smalle’ betekenis van opbrengstgericht werken, zoals beschreven in het integraal eindrapport van Streef; Doolaard, 2013). Opbrengstgericht werken is een manier van denken

en handelen, waarbij leerkrachten zich bewust zijn van het leerdoel dat zij willen bereiken, de aanpak die hiervoor gehanteerd wordt en de huidige stand van zaken in relatie tot dat doel. De werkwijze zou omschreven kunnen worden als de bekende Plan-Do-Check-Act-cyclus in combinatie met een samenwerkingsgericht team en het stellen van hoge, uitdagende doelen. In het kader van de BOPO onderzoekslijn ‘het streven naar kwaliteit in scholen voor primair onderwijs’ wordt met verschillende schoolverbetertrajecten (zogenaamde ‘Streef deelprojecten’, zie ook het integraal eindrapport; Doolaard, 2013) ingegaan op opbrengstgericht werken, met daarbinnen het stellen van hoge doelen en het streven naar excellentie. Het effect van dergelijke ondersteuning en de algemene implementatie van nieuw onderwijsbeleid hangt samen met de mate waarin de nieuwe onderwijspraktijken door de leerkrachten worden geaccepteerd en gebruikt (Hattie, 2003; Hill, 2007; Scheerens & Bosker, 1997; Stipek, Givvin, Salmon, & MacGyvers, 2001). Om deze attitude en het handelen van leerkrachten in kaart te brengen en om de opbrengstgerichtheid van schoolteams te kunnen inschatten, is twee keer een uitgebreide, digitale vragenlijst afgenomen bij alle leerkrachten van alle scholen die deelnamen aan Streef. De vragenlijst was onderdeel van de flankerende dataverzameling voor alle afzonderlijke Streef-interventies. De vragenlijst beslaat verschillende thema’s rond onderwijskwaliteit en gaat in op handelen en attitudes met betrekking tot opbrengstgericht werken, toetsgebruik en –analyse, nadruk op basissvaardigheden en referentieniveaus, differentiatie en samenwerking en professionalisering. De eerste afname vond plaats in het voorjaar 2010. Dezelfde vragenlijst is nogmaals afgenomen zomer 2012, na afronding van de verschillende schoolverbetertrajecten. Voor de theoretische onderbouwing van de vragenlijst wordt verwezen naar de deelrapportage over de eerste afname (Deunk, Doolaard & Hofman, 2011).

In de huidige rapportage wordt met name ingegaan op de antwoorden van de respondenten die de vragenlijst zowel in 2010 als in 2012 hebben ingevuld. Daarnaast wordt gekeken naar de antwoorden van schoolleiders en intern begeleiders en naar de kenmerken van leerkrachtenteams. De volgende onderzoeksvragen zullen worden beantwoord:

- Wat is de attitude en het handelen van leerkrachten met betrekking tot aspecten van opbrengstgericht werken, te weten prestatieniveaus en doelen, referentieniveaus en toetsgebruik en -analyse?
- Wat is de attitude en het handelen van leerkrachten met betrekking tot differentiatie?
- Wat is de attitude en het handelen van leerkrachten met betrekking tot professionalisering en samenwerking in het team?
- Wat is de attitude en het handelen van schoolleiders en intern begeleiders en hoe verhoudt dit zich tot de attitude en het handelen van leerkrachten?
- Wat is de samenhang tussen de opbrengstgerichtheid van leerkrachtenteams en de leerwinst van hun leerlingen?

Hierbij zal een vergelijking worden gemaakt tussen leerkrachten die deel hebben genomen aan een deelproject van Streef en leerkrachten die niet zelf actief hebben deelgenomen. De resultaten van 2010 en 2012 zullen gepresenteerd worden, met nadruk op veranderingen in de tijd en de meest recente stand van zaken.

2 Methode

2.1 Respondenten

In totaal hebben er 120 scholen deelgenomen aan Streef. Scholen participeerden in principe alleen met de onder- of de bovenbouw in een van de interventies, waardoor de andere bouw fungeert als controlegroep voor een van de andere interventies. Voor een uitgebreide beschrijving van de werving wordt verwezen naar het integraal eindrapport van het Streef project (Doolaard, 2013). Alle leerkrachten van alle scholen werd gevraagd de Streef vragenlijst in te vullen, of zij nu zelf actief deelnamen aan een interventie of niet. De eerste afname vond in het voorjaar 2010 plaats, voor de start van de interventies; de tweede afname vond plaats aan het eind van het schooljaar 2011-2012. Voor de leerkrachten die deelnamen aan een interventie in het schooljaar 2010-2011 viel de tweede afname van de vragenlijst dus niet direct na afloop van de interventie, maar een schooljaar later.

Voorjaar 2010 hebben 793 respondenten afkomstig van 92 scholen de vragenlijst ingevuld¹ (een respons van 79%). Zomer 2012 hebben 641 respondenten afkomstig van 105 scholen de vragenlijst ingevuld (514 leerkrachten, 60 schoolleiders en 60 intern begeleiders; een respons van 55%²). 378 deelnemers, afkomstig van 79 scholen, hebben de vragenlijst beide keren ingevuld. Deze groep bestaat uit 15 schoolleiders, 18 intern begeleiders, 338 leerkrachten en 7 schoolleiders of intern begeleiders die structureel minimaal 4 dagdelen voor de klas staan. In deze rapportage staan de lesgevende respondenten centraal, dus de leerkrachten en de schoolleiders en intern begeleiders met lesgevende taken (vanaf nu allemaal ‘leerkrachten’ genoemd). Deze groep van 345 leerkrachten (ook afkomstig van 79 scholen) geeft ten tijde van 2012 gemiddeld bijna 20 jaar les, variërend van 0 jaar tot 42 jaar. Ongeveer de helft geeft in 2012 les in de onderbouw (171 leerkrachten, 51%). De helft van de leerkrachten (166, 49%) heeft daarbij een combinatiegroep, meestal een combinatie van groep 1 en 2 (65 leerkrachten, 19%), gevolgd door de combinatie van groep 3 en 4 (25 leerkrachten, 7%), groep 7 en 8 (23 leerkrachten, 7%) en groep 5 en 6 (22 leerkrachten, 6%).

Van de 345 leerkrachten heeft iets minder dan de helft (163 leerkrachten, 47%) actief deelgenomen aan een van de deelprojecten van Streef. 23 leerkrachten namen deel aan een project over referentieniveaus, 85 aan een project over opbrengstgericht werken, 25 aan een

¹ Er is een vergelijking gemaakt tussen 2010 en 2012. Er is een extra voormeting gedaan bij een kleine groep leerkrachten in 2011, maar deze wordt in de huidige rapportage buiten beschouwing gelaten.

² Een respons van 55% lijkt laag, maar is dat niet voor een nameting. De meeste scholen hebben met een deel van hun team deelgenomen aan een van de deelprojecten van Streef in het schooljaar '10-'11. Het is logisch dat de respons een jaar, aan het eind van het schooljaar '11-'12, wat af is genomen. Het aantal scholen waar de respondenten werkzaam zijn is bovendien groot (N=105), waardoor de representativiteit van de respondentengroep gewaarborgd is.

project over excellentie en 44 aan een project over leren lezen³. De leerkrachten die hebben deelgenomen aan een Streef-project en degenen die niet hebben deelgenomen lijken op elkaar wat betreft aantal jaren leservaring en percentage leerkrachten dat lesgeeft aan een combinatiegroep. Er is wel een verschil in de bouw waarin de leerkrachten lesgeven: bijna twee derde (64%) van de Streefdeelnemers geeft les in de bovenbouw. Bij de bespreking van de resultaten zal onderscheid gemaakt worden tussen leerkrachten die wel en niet aan een van de deelprojecten van Streef hebben deelgenomen.

Deze rapportage richt zich op de 345 leerkrachten die de vragenlijst zowel in 2010 als in 2012 hebben ingevuld. In beide afnamejaren is de vragenlijst echter door een grotere groep respondenten ingevuld. In totaal hebben 760 leerkrachten de vragenlijst in 2010 ingevuld, waarvan 415 leerkrachten alleen in 2010 hebben geantwoord. De twee groepen verschillen iets in leservaring en deelname aan de Streefprojecten. In 2010 hadden de respondenten die beide keren de vragenlijst hebben ingevuld bijna 18 jaar leservaring. Deze groep leerkrachten heeft gemiddeld bijna 2 jaar meer ervaring dan de groep leerkrachten die de vragenlijst alleen in 2010 heeft ingevuld ($t(755)=-2.26$, $p=0.024$; zie ook de bijlage). In 2012 is gevraagd naar *actieve* deelname, dat wil zeggen dat leerkrachten zelf hebben deelgenomen aan de Streef trainingen. Over 2010 is er alleen informatie over *algemene* deelname, dat wil zeggen dat de school en bouw van de betreffende leerkracht deelnam aan een van de Streef projecten. Het kan zijn dat bijvoorbeeld een duo-partner of een collega uit een parallelklas niet deelnam aan de training. De aantallen respondenten met algemene deelname zijn vanzelfsprekend hoger dan de aantallen respondenten met actieve deelname. De groep leerkrachten die beide vragenlijsten heeft ingevuld kan alleen wat betreft *algemene* deelname vergeleken worden met de groep leerkrachten die de vragenlijst alleen in 2010 heeft ingevuld. Onder de groep die beide vragenlijsten heeft ingevuld bevinden zich meer leerkrachten die hebben deelgenomen aan een Streef project dan onder de groep die alleen in 2010 heeft geantwoord (respectievelijk 77% en 67%; $\chi^2(1)=10.1$, $p=0.002$; zie ook de bijlage). De geselecteerde groep van 345 leerkrachten die de vragenlijsten in 2010 en 2012 heeft ingevuld verschilt dus iets in samenstelling van de totale groep leerkrachten die alleen in 2010 heeft ingevuld, maar over het algemeen verschillen zij niet in hun antwoordpatronen. Door middel van t-toetsen en χ^2 -toetsen is gecontroleerd of de antwoorden van de twee groepen significant van elkaar verschillen. De enkele keer dat er verschillen optreden, zullen deze worden gerapporteerd in voetnoten bij de betreffende paragraaf en in de bijlage.

Tenslotte zijn er in 2012 specifiek vragen aan de schoolleiders en intern begeleiders gesteld. In totaal hebben 60 schoolleiders en 60 intern begeleiders, afkomstig van 75 scholen, de vragen beantwoord. Hun antwoorden zullen apart behandeld worden aan het eind van het

³ Zeven leerkrachten hebben actief aan twee deelprojecten deelgenomen.

rapport. Waar relevant zullen de antwoorden worden vergeleken met die van de leerkrachten die de vragenlijst hebben ingevuld in 2012 en die actief hebben deelgenomen aan een van de Streefprojecten (N=262). Er is voor gekozen om alleen de leerkrachten te selecteren die daadwerkelijk hebben deelgenomen aan een van de Streefprojecten om een betere vergelijking te kunnen maken. De schoolleiders en IB-ers zijn immers automatisch al meer betrokken bij de interventie doordat zij in eerste instantie hebben ingestemd met deelname aan het onderzoek met hun school.

2.2 Analyseplan

Bij elk deelonderwerp in de vragenlijst worden twee verschillende vragen gesteld:

1. Zijn leerkrachten veranderd in hun attitude, verwachtingen en het handelen in de periode 2010-2012 en wat is de stand van zaken in 2012?
2. Hangen deze attitude, verwachtingen en het handelen en verandering hierin samen met deelname aan een van de deelprojecten van Streef?

Voor elke set vragen, index of schaal wordt in tabellen weergegeven hoe de groep leerkrachten in de controlegroep en in de Streef-groep antwoordden in 2010 en in 2012 en of beide groepen op de twee tijdstippen van elkaar verschilden. Daarnaast wordt per groep de mate van verandering over de tijd weergegeven. De resultaten worden descriptief toegelicht en samengevat. Waar relevant worden statistische analyses toegepast. In elke tabel wordt aangegeven of de controlegroep en de Streef-groep van elkaar verschillen in 2010 en 2012 (in de tabellen aangegeven met respectievelijk de noten ^a en ^b), of er een verandering is van 2010 tot 2012 (in de tabellen aangegeven met de noot ^c) en of de groepen van elkaar verschillen in mate van verandering (in de tabellen aangegeven met de noot ^d). Hierbij wordt een significantieniveau van $p < 0.05$ gehanteerd. De belangrijkste verschillen en opmerkelijkste resultaten worden daarnaast in detail in de tekst toegelicht.

De statistische analysemethodes die gebruikt worden hangen samen met de antwoordopties bij elke vraag, index of schaal. Bij vragen met een oplopende antwoordschaal op interval- of ratio-niveau, zoals stellingen met antwoordopties lopend van helemaal eens tot helemaal oneens of vragen naar aantallen leerlingen, worden verschillen tussen de groepen per meetmoment geanalyseerd met t-toetsen. Veranderingen over de tijd worden per groep bekeken met gepaarde t-toetsen. Daarnaast wordt per leerkracht het verschil tussen het antwoord op de voormeting en op de nameting uitgerekend. De gemiddelde toe- of afname van de leerkrachten in de controlegroep en de Streef-groep wordt vertaald in mate van aanpassing relatief aan de antwoordschaal en vervolgens vergeleken met een t-toets. Bij vragen op ordinaal niveau worden de groepen op beide meetmomenten vergeleken door een kruistabel te maken en een Chi²-toets toe te passen. Vervolgens wordt bepaald of de leerkrachten in van 2010 tot 2012 in positieve of negatieve zin zijn veranderd, zonder daarbij uitspraken te doen over de grootte van deze individuele veranderingen. Bij dit type vragen

wordt steeds het aantal leerkrachten genoemd dat ‘achteruit’ en ‘vooruit’ is gegaan in de periode 2010-2012. Per onderzoeksgroep wordt met een tekentoets gekeken of de verandering in het algemeen positief of negatief was. Verschillen in verandering tussen de groepen worden in kaart gebracht met een kruistabel en bijbehorende Chi^2 -toets.

Daarna is gekeken naar de attitude en verwachtingen van schoolleiders en intern begeleiders en de mate waarin zij andere opvattingen en handelwijzen hebben dan de leerkrachten. Dit betreft alleen data over 2012, dus er kan niet gekeken worden naar veranderingen over de tijd. Er wordt gebruik gemaakt van beschrijvende statistiek. Daarnaast worden de groep schoolleiders en IB-ers en de leerkrachten vergeleken. Verschillen worden in kaart gebracht met Chi^2 -toetsen en t-toetsen, afhankelijk van het meetniveau van de variabelen (in de betreffende tabellen aangegeven met noot ^a). Zoals in de eerdere vergelijkingen wordt een significantieniveau van $p < 0.05$ gehanteerd.

De vraag naar de samenhang tussen de opbrengstgerichtheid van leerkrachtenteams en de leerwinst van leerlingen tenslotte wordt beantwoord door op 9 belangrijke kenmerken van opbrengstgericht werken een clusteranalyse uit te voeren. Dit leidt tot vier clusters, of typen leerkrachtenteams. Vervolgens zijn de “opbrengstgerichte” leerkrachtenteams afgezet tegen de drie andere typen teams en is gekeken hoe het type team samenhangt met de leerwinst van de leerlingen op verschillende vakken. De effecten van de opbrengstgerichtheid van teams zijn geschat met meerniveau-analyses (scholen en leerlingen). Er is gecontroleerd voor relevante pre-tests, zoals de term ‘leerwinst’ al aangeeft.

3 Resultaten

3.1 Opbrengstgericht werken

3.1.1 Eindniveaus en prestaties

Bij opbrengstgericht werken voor de basisvakken taal en rekenen gaat het erom dat leerkrachten van te voren een prestatiedoel hebben gesteld en dat zij weten hoe hun leerlingen presteren ten opzichte van dit doel. Daarnaast is het belang dat leerkrachten hechten aan opbrengstgericht werken van invloed op de mate van opbrengstgericht werken.

De leerkrachten is eerst gevraagd of zij tevreden zijn met het huidige prestatieniveau van hun leerlingen (zie tabel 3.1).

Tabel 3.1 Tevredenheid huidige prestatieniveau van de leerlingen
“Vindt u het niveau van uw leerlingen op dit moment goed genoeg bij...”

	2010					2012					Verandering		
	N	Ja	Nee 1	Nee 2	Nee 3	N	Ja	Nee 1	Nee 2	Nee 3	N	-	+
(Aanvankelijk) rekenen													
Controle ^c	169	62	93	11	3	172	80	85	4	3	162	24	42
		37%	55%	7%	2%		47%	49%	2%	2%		15%	26%
Streef ^c	154	39	99	14	2	158	61	86	10	1	150	24	48
		25%	63%	9%	1%		39%	54%	6%	1%		16%	32%
Taal (groep 1&2)													
Controle	57	19	34	4	0	119	33	28	2	0	51	8	15
		33%	60%	7%	0%		52%	44%	3%	0%		35%	29%
Streef	28	8	20	0	0	25	11	14	0	0	23	3	8
		29%	71%	0%	0%		44%	56%	0%	0%		13%	35%
Technisch lezen													
Controle	68	14	48	6	0	81	22	51	8	0	53	7	13
		21%	71%	9%	0%		27%	63%	10%	0%		13%	25%
Streef ^c	64	6	47	11	0	75	14	54	6	1	51	3	13
		9%	73%	17%	0%		19%	72%	8%	1%		6%	26%
Begrijpend lezen													
Controle	90	13	57	19	1	90	16	56	14	4	73	14	17
		14%	63%	21%	1%		18%	62%	16%	4%		20%	23%
Streef	108	6	73	23	6	117	15	64	30	8	98	19	22
		6%	68%	21%	6%		13%	55%	26%	7%		19%	22%

1= Nee, enkele leerlingen zouden beter moeten presteren; 2= Nee, zeker de helft zou beter moeten presteren; 3= Nee, bijna de hele klas zou beter moeten presteren; -= minder tevreden; += meer tevreden

^c= significant meer leerkrachten veranderen in positieve/negatieve richting (p<0.05)

Zowel in 2010 als in 2012 zijn de leerkrachten in de Streef-groep iets minder tevreden over het niveau van de leerlingen dan de leerkrachten in de controlegroep, hoewel de groepen niet significant van elkaar verschillen. In het algemeen zijn de leerkrachten het minst tevreden

over het begrijpend leesniveau van hun leerlingen: slechts 6%-14% van de leerkrachten in 2010 en 13%-18% in 2012 is tevreden met het prestatieniveau van de leerlingen en 20%-33% van de leerkrachten vindt dat de prestaties voor begrijpend lezen bij de halve tot de hele klas achterblijven.

Als gekeken wordt naar verandering in tevredenheid van 2010 tot 2012, valt op dat ongeveer de helft van de leerkrachten even (on)tevreden is in 2012 als in 2010⁴. Voor het domein rekenen zijn in beide groepen relatief veel leerkrachten meer tevreden geworden (Controle: $p=0.036$; Streef: $p=0.006$). Op het gebied van technisch lezen zijn binnen de Streef-groep relatief veel leerkrachten meer tevreden geworden in de loop der tijd ($p=0.021$).

De tevredenheid van de leerkracht over het prestatieniveau van de klas betreft de prestaties van de leerlingen aan het eind van het schooljaar en is daarmee een soort evaluatie achteraf. Binnen opbrengstgericht werken is het belangrijk dat vooraf al duidelijk wordt gemaakt naar welk niveau gestreefd wordt. Er is daarom aan de leerkrachten gevraagd of zij voor ogen hebben wat het eindniveau is dat zij nastreven met de leerlingen, en zo ja op welke manier dat niveau is geëxpliciteerd. De verschillende manieren waarop leerkrachten niveaus kunnen expliciteren zijn samengevat in drie categorieën⁵, namelijk: de leerkracht heeft geen beeld van de gewenste eindniveaus, de leerkracht heeft de eindniveaus individueel voor zichzelf vastgesteld, of de leerkracht heeft de eindniveaus in overleg met de intern begeleider of het team vastgesteld (tabel 3.2).

Voor de domeinen rekenen, taal en technisch lezen hebben de meeste leerkrachten wel einddoelen gesteld. Opvallend is dat een relatief grote groep van ongeveer 15% van de leerkrachten in 2010 en 14% tot 27% van de leerkrachten in 2012 niet voor ogen heeft welk begrijpend leesniveau bereikt moet zijn aan het eind van het schooljaar. Een derde tot de helft van de leerkrachten verandert tussen 2010 en 2012 in de manier waarop doelen worden vastgelegd. Sommige leerkrachten gaan meer expliciteren (door doelen te stellen terwijl ze dat eerst helemaal niet deden of door in overleg doelen te stellen terwijl ze dat eerst zelf deden) maar sommigen juist ook minder. Vooral voor de rekendoelen in de Streef-groep is de verandering positief: relatief veel leerkrachten gaan de rekendoelen meer expliciteren in de loop der tijd en relatief weinig leerkrachten gaan minder expliciteren ($p=0.000$). Hierin verschillen deze leerkrachten van de leerkrachten in de controlegroep, waarin ongeveer evenveel leerkrachten meer en minder zijn gaan expliciteren ($\chi^2(2)=6.6$, $p=0.037$). In 2010 stelden de leerkrachten binnen de Streef-groep hun rekendoelen nog meestal zelf, terwijl van

⁴ Af te leiden uit tabel 1: het percentage leerkrachten dat niet veranderd is qua tevredenheid is 100% min de percentages positieve en negatieve veranderingen in de kolom 'verandering'.

⁵ Leerkrachten konden via een open vraag ook aangeven dat ze de doelen op een andere manier vastleggen. Deze antwoordoptie is hier buiten beschouwing gelaten.

de leerkrachten in de controlegroep twee derde de doelen al in overleg vaststelde⁶. In 2012 is dit verschil opgeheven en stelt ruim 40% van de leerkrachten in beide groepen de gewenste eindniveaus voor rekenen in overleg met de IB-er of het team vast.

Tabel 3.2 Wijze waarop men het gewenste eindniveau heeft geëxpliciteerd

“Weet u wat het niveau is dat u met de leerlingen wilt behalen aan het eind van het schooljaar voor...”

	2010				2012				Verandering		
	N	Geen	Zelf	Samen	N	Geen	Zelf	Samen	N	-	+
(Aanvankelijk) rekenen^{bd}											
Controle	145	4	91	50	148	10	75	63	122	19	22
		3%	63%	35%		7%	51%	43%		16%	18%
Streef ^c	129	3	98	28	139	1	79	59	114	7	30
		2%	76%	22%		1%	57%	42%		6%	26%
Taal (groep 1&2)											
Controle	38	1	16	21	48	1	15	32	25	2	7
		3%	42%	55%		2%	31%	67%		8%	28%
Streef	17	0	8	9	19	0	7	12	10	0	3
		0%	47%	53%		0%	37%	63%		0%	30%
Technisch lezen											
Controle	63	1	29	33	70	5	36	29	43	13	7
		2%	46%	52%		7%	51%	41%		30%	16%
Streef	58	2	31	25	60	0	35	25	38	9	9
		3%	53%	43%		0%	58%	42%		24%	24%
Begrijpend lezen											
Controle	84	11	55	18	85	23	43	19	67	16	10
		13%	66%	21%		27%	51%	22%		24%	15%
Streef	97	15	57	25	106	15	59	32	81	16	20
		16%	59%	26%		14%	56%	30%		20%	25%

geen= geen beeld van gewenste eindniveaus, zelf= eindniveaus voor zichzelf vastgesteld, samen= eindniveaus vastgesteld in overleg met IB-er of team; -= minder expliciteren; += meer expliciteren

^b= controlegroep en Streef-groep verschillen significant in 2012 ($p < 0.05$)

^c= significant meer leerkrachten veranderen in positieve/negatieve richting ($p < 0.05$)

^d= controlegroep en Streef-groep verschillen significant in mate van verandering ($p < 0.05$)

De mate waarin een leerkracht opbrengstgericht zal werken hangt samen met hoe hij of zij hier tegenover staat. In 2010 en 2012 zijn vier verschillende stellingen voorgelegd die hierop ingaan. Deze stellingen zijn samengevoegd in de index “Attitude Opbrengsten” (tabel 3.3). De index heeft een schaal van 1 (negatieve attitude) tot 4 (positieve attitude), waarbij antwoorden rond de 2,5 een neutrale houding weergeven. De leerkrachten in de controlegroep zijn zowel in 2010 als in 2012 redelijk neutraal in hun attitude ten opzichte van opbrengsten. De

⁶ Voor aanvankelijk rekenen is er een verschil tussen de groep leerkrachten die de vragenlijst in beide jaren heeft ingevuld en de groep die de vragenlijst alleen in 2010 heeft ingevuld. De groep die beide jaren heeft ingevuld stelt de doelen vaker samen (29% vs 20%) en de groep die alleen geantwoord heeft in 2010 stelt de doelen vaker zelf (77% vs 69%; $\chi^2(2)=6.43$, $p=0.040$; zie ook de bijlage).

leerkrachten in de Streef-groep krijgen in de loop der jaren een positievere attitude en veranderen van neutraal naar redelijk positief, met een verschuiving van ruim 7,5% op de antwoordschaal ($t(148)=-2.7$, $p=0.007$). Deze verandering leidt ertoe dat de leerkrachten in de Streef-groep in 2012 positiever tegenover opbrengsten staan dan de leerkrachten in de controlegroep ($t(315)=-2.3$, $p=0.021$).

Tabel 3.3 Index attitude opbrengsten

	2010			2012 ^b			Verandering		
	N	Gem*	Sd	N	Gem	Sd	N	Gem	Sd
Controle	171	2,57	0,91	163	2,67	0,95	155	+0,07 (2,3%)	1,00
Streef ^c	155	2,66	0,94	154	2,90	0,84	149	+0,23 (7,7%)	1,06

*Op een schaal van 1 (negatieve attitude) tot 4 (positieve attitude)

^b= controlegroep en Streef-groep verschillen significant in 2012 ($p<0.05$)

^c= significante verandering van 2010 naar 2012 ($p<0.05$)

Vervolgens is specifieker ingegaan op de nadruk op basisvaardigheden taal en rekenen (tabel 3.4). Drie verschillende vragen over het belang van nadruk op taal en rekenen zijn samengenomen in een schaal Basisvaardigheden⁷. Deze schaal loopt van 1 (helemaal oneens met de nadruk op basisvaardigheden) tot 5 (helemaal eens), waarbij antwoorden rond 3 een neutrale houding weergeven. De leerkrachten in beide groepen zijn het zowel in 2010 als in 2012 licht eens met de nadruk op basisvaardigheden. De leerkrachten in de Streef-groep worden het in de periode 2010-2012 iets nog meer eens ($t(148)=-2.0$, $p=0.042$).

Tabel 3.4 Schaal basisvaardigheden

	2010			2012			Verandering		
	N	Gem*	Sd	N	Gem	Sd	N	Gem	Sd
Controle	171	3,52	0,67	160	3,59	0,70	152	+0,04 (1,0%)	0,68
Streef ^c	155	3,51	0,68	154	3,65	0,73	149	+0,14 (3,5%)	0,81

*Op een schaal van 1 (helemaal oneens) tot 5 (helemaal eens)

^c= significante verandering van 2010 naar 2012 ($p<0.05$)

3.1.2 Referentieniveaus

De nadruk op basisvaardigheden sluit aan bij het belang dat de overheid hecht aan taal en rekenen met de invoering van landelijke referentieniveaus. Om te beginnen is leerkrachten gevraagd of zij bekend zijn met de referentieniveaus en zo ja, hoe zij die kennis hebben opgedaan. Deze vragen zijn samengevat in de index Bekendheid Referentieniveaus (tabel 3.5). In 2010 wist nog ruim de helft van de leerkrachten niet wat de referentieniveaus waren. Door aandacht in de media en door de Streef-projecten zelf is te verwachten dat de bekendheid met referentieniveaus over de tijd stijgt.

⁷ De schaal Basisvaardigheden is betrouwbaar met in de voormeting $\alpha=0,71$ en in de nameting: $\alpha=0,77$.

In 2012 is de bekendheid met referentieniveaus inderdaad toegenomen. Zowel in de controlegroep als in de Streef-groep is het aantal leerkrachten dat nooit van referentieniveaus heeft gehoord gedaald naar respectievelijk 37% en 21% en is het aantal leerkrachten dat er op school over spreekt gestegen naar 27% tot 29% (Controle: $p=0.00$; Streef: $p=0.00$). Hoewel beide groepen bekender zijn geworden met de referentieniveaus, is de Streef-groep meer veranderd dan de controlegroep ($\chi^2(2)=12.6$, $p=0.002$). De leerkrachten in de Streef-groep zijn in 2012 daardoor beter bekend met de referentieniveaus dan de leerkrachten in de controlegroep ($\chi^2(3)=39.4$, $p=0.00$).

Tabel 3.5: Index bekendheid referentieniveaus

	2010					2012 ^b					Verandering ^d		
	N	1	2	3	4	N	1	2	3	4	N	-	+
Controle ^c	169	92	15	36	26	169	62	23	39	45	159	22	57
		54%	9%	21%	15%		37%	14%	23%	27%		14%	36%
Streef ^c	156	78	10	45	23	157	19	60	33	45	151	30	75
		50%	6%	29%	15%		21%	38%	21%	29%		20%	50%

1= niet van gehoord, 2= wel eens van gehoord, 3= over gelezen of in verdiept, 4= over gesproken op school,

-= bekendheid afgenomen += bekendheid toegenomen.

^b= controlegroep en Streef-groep verschillen significant in 2012 ($p<0.05$)

^c= significant meer leerkrachten veranderen in positieve/negatieve richting ($p<0.05$)

^d= controlegroep en Streef-groep verschillen significant in mate van verandering ($p<0.05$)

Alvorens verdere vragen over referentieniveaus te stellen is een korte uitleg gegeven, om iedereen dezelfde basisinformatie over referentieniveaus te verschaffen. Daarna zijn leerkrachten verschillende uitspraken voorgelegd over het nut en de bruikbaarheid van referentieniveaus, om hun attitude in kaart te kunnen brengen.

Het blijkt dat maar erg weinig leerkrachten een negatief effect verwachten van de referentieniveaus (tabel 3.6). In 2010 verwachten de meeste leerkrachten een positief effect van referentieniveaus op hun dagelijks werk en de prestaties van de leerlingen. De leerkrachten in de Streef-groep zijn hier significant positiever over dan de leerkrachten in de controlegroep (dagelijks werk: $\chi^2(2)=14.8$, $p=0.001$; prestaties leerlingen: $\chi^2(2)=9.2$, $p=0.010$). In 2012 is de positieve verwachting echter afgenomen en denken leerkrachten vaker dat de referentieniveaus geen invloed zullen hebben, of weten ze het niet. Deze daling in positivisme geldt voor zowel de leerkrachten in de controlegroep als in de Streef-groep voor invloed op het dagelijks werk (Controle: $p=0.015$; Streef: $p=0.000$). Wat betreft de invloed van de referentieniveaus op de prestaties van de leerlingen zijn vooral veel leerkrachten in de Streef-groep minder overtuigd positief geworden ($p=0.00$).

Tabel 3.6 Invloed referentieniveaus
“In welke mate zijn de referentieniveaus van invloed op...”

	2010				2012				Verandering		
	N	Neg	Geen	Pos	N	Neg	Geen	Pos	N	-	+
Uw dagelijks werk^a											
Controle ^c	169	7	72	90	164	2	99	63	155	45	24
		4%	43%	53%		1%	60%	38%		30%	16%
Streef ^c	155	2	39	114	155	2	79	74	150	51	11
		1%	25%	74%		1%	51%	48%		34%	7%
De prestaties van leerlingen^a											
Controle	171	6	75	90	163	1	92	70	155	43	28
		4%	44%	53%		1%	56%	43%		28%	18%
Streef ^c	154	1	49	104	154	1	83	70	147	51	18
		1%	32%	68%		1%	54%	46%		35%	12%

Neg= negatieve invloed; geen= geen invloed of weet niet; pos= positieve invloed;

-= meer negatieve verwachting, += meer positieve verwachting

^a= controlegroep en Streef-groep verschillen significant in 2010 ($p < 0.05$)

^c= significant meer leerkrachten veranderen in positieve/negatieve richting ($p < 0.05$)

De attitude van leerkrachten tegenover referentieniveaus is gemeten met 9 stellingen, waarvan de laatste 4 (stelling f t/m i) alleen in 2012 zijn voorgelegd. Leerkrachten konden op een schaal van 1 (helemaal oneens) tot 5 (helemaal eens) aangeven in hoeverre de stellingen op hen van toepassing waren (tabel 3.7). Over het algemeen blijken leerkrachten gematigd positief. Referentieniveaus kunnen volgens de leerkrachten met name een rol spelen bij het verduidelijken van wat van leerlingen verwacht kan worden en bij het differentiëren (stelling g en h), hoewel de leerkrachten ook aangeven dat de meeste leerkrachten al wel weten wat de leerdoelen zijn (stelling a).

De groepen leerkrachten verschillen van mening over stelling d: *“Ik vind het stellen van landelijke referentieniveaus oneerlijk voor leerkrachten die toevallig een klas met veel zwakke leerlingen hebben”*. Leerkrachten in de Streef-groep zijn het hier zowel in 2010 als 2012 minder mee eens dan hun collega's in de controlegroep (2010: $t(325)=2.09$, $p=0.038$; 2012: $t(312)=2.15$, $p=0.033$) en geven hiermee een positievere houding tegenover referentieniveaus weer. De leerkrachten in beide groepen blijven voor het grootste gedeelte in gelijke mate instemmen met de stellingen in de periode 2010-2012. Alleen voor de stelling a: *“de meeste leerkrachten weten wel wat de leerdoelen zijn en hoe ze daar aan moeten werken”* is er een verschil. Leerkrachten uit de Streef-groep stemmen hier in 2012 meer mee in dan in 2010 ($t(147)=-3.12$, $p=0.002$) en hun attitudeverandering is significant groter dan die in de controlegroep ($t(299)=-1.99$, $p=0.047$).

Tabel 3.7 Attitude referentieniveaus

“Hieronder volgt een aantal uitspraken die te maken hebben met de referentieniveaus en het gebruik daarvan in de klas. Kunt u aangeven in hoeverre de uitspraken op u van toepassing zijn?”

	2010			2012			Verandering		
	N	Gem	Sd	N	Gem	Sd	N	Gem	Sd
a. De meeste leerkrachten weten wel wat de leerdoelen zijn en hoe ze daar aan moeten werken.^d									
Controle	171	3,68	0,73	161	3,69	0,71	153	+0,03 (1%)	0,83
Streef ^c	157	3,55	0,85	153	3,73	0,63	148	+0,22 (6%)	0,83
b. Ik denk dat het vooral voor nieuwe, onervaren leerkrachten goed is als er duidelijke referentieniveaus komen.									
Controle	170	3,72	0,93	161	3,63	0,86	152	-0,11 (3%)	1,05
Streef	154	3,58	1,06	154	3,58	0,93	147	+0,01 (0%)	1,15
c. Ik vind het stellen van landelijke referentieniveaus oneerlijk voor scholen met veel zwakke leerlingen.									
Controle	167	3,11	0,93	161	3,04	0,82	149	-0,07 (2%)	1,00
Streef	157	2,92	0,92	153	2,87	0,96	148	-0,04 (1%)	1,07
d. Ik vind het stellen van landelijke referentieniveaus oneerlijk voor leerkrachten die toevallig een klas met veel zwakke leerlingen hebben.^{ab}									
Controle	171	3,05	0,89	161	2,98	0,87	153	-0,05 (1%)	0,98
Streef	156	2,84	0,90	153	2,76	0,92	147	-0,05 (1%)	1,01
e. Ik vind het voordeel van referentieniveaus dat de leerkracht gestimuleerd wordt ook de zwakke leerlingen op een minimumniveau te brengen.									
Controle	170	3,72	0,75	160	3,62	0,79	151	-0,12 (3%)	0,87
Streef	157	3,78	0,78	154	3,75	0,79	149	-0,04 (1%)	0,93
f. Ik denk dat het moeilijk wordt voor leerkrachten om te voldoen aan de referentieniveaus.									
Controle				161	3,12	0,80			
Streef				153	2,90	0,84			
g. Ik vind het voordeel van referentieniveaus dat zij duidelijk maken wat er van leerlingen verwacht kan worden.									
Controle				161	3,84	0,65			
Streef				153	3,95	0,66			
h. Ik denk dat referentieniveaus leerkrachten kunnen helpen om beter te differentiëren tussen leerlingen.									
Controle				161	3,68	0,74			
Streef				152	3,79	0,79			
i. Ik maak gebruik van referentieniveaus bij het evalueren van toetsresultaten.									
Controle				160	2,86	0,97			
Streef				154	2,81	1,02			

1= helemaal oneens, 2= oneens, 3= niet oneens/eens, 4= eens, 5= helemaal eens

^a= controlegroep en Streef-groep verschillen significant in 2010 ($p < 0.05$)

^b= controlegroep en Streef-groep verschillen significant in 2012 ($p < 0.05$)

^c= significante verandering van 2010 naar 2012 ($p < 0.05$)

^d= controlegroep en Streef-groep verschillen significant in mate van verandering ($p < 0.05$)

3.1.3 Toetsgebruik en rapportage

Toetsen zijn erg belangrijk binnen het opbrengstgericht werken. Toetsresultaten geven namelijk belangrijke informatie over het prestatieniveau van de leerling en in hoeverre hij of zij nog van het gestelde doel verwijderd is. Met name gestandaardiseerde, methodeoverstijgende toetsen zijn belangrijk voor een objectief beeld van de prestaties van individuele leerlingen en van de klas als geheel. Het overgrote deel van de deelnemende leerkrachten maakt gebruik van gestandaardiseerde toetsen uit het Cito leerlingvolgsysteem (tabel 3.8). Leerkrachten geven daarnaast aan de methodeoverstijgende toetsen voor het verkrijgen van informatie over leerlingen relatief belangrijk te vinden, in vergelijking met de informatiewaarde van de eigen kennis en indruk, observaties in de klas, dagelijks schriftelijk werk en methodetoetsen (tabel 3.9). In 2010 waren de leerkrachten in de controlegroep positiever over de informatiewaarde van methodeoverstijgende toetsen dan de leerkrachten in de Streef-groep ($t(292)=2.25$, $p=0.025$), in 2012 is dit verschil verdwenen.

Tabel 3.8: Gebruik methodeoverstijgende toetsen

“Maakt u in uw klas gebruik van methodeoverstijgende toetsen?”

	2010				2012			
	N	Geen	Cito LOVS	Anders	N	Geen	Cito LOVS	Anders
Controle	175	10 (6%)	160 (91%)	5 (3%)	175	10 (6%)	159 (91%)	6 (3%)
Streef	157	5 (3%)	145 (92%)	7 (5%)	159	7 (4%)	151 (95%)	1 (1%)

Tabel 3.9: Index waarde toets

“Hoe informatief zijn methodeoverstijgende toetsen voor het verkrijgen van tussentijdse informatie over leerlingen”

	2010 ^a			2012			Verandering		
	N	Gem*	Sd	N	Gem	Sd	N	Gem	Sd
Controle	154	3,36	0,97	142	3,32	1,04	128	-0,08 (2,7%)	1,36
Streef	140	3,08	1,16	138	3,22	1,15	122	+0,15 (5%)	1,38

*Op een schaal van 1 (minst informatief) tot 4 (meest informatief)

^a= controlegroep en Streef-groep verschillen significant in 2010 ($p<0.05$)

Om toetsresultaten goed te kunnen analyseren en interpreteren kunnen verschillende rapporten worden gemaakt. Globaal gezien kan er gekozen worden voor rapportages op leerlingniveau en rapportages op groepsniveau. Daarnaast kan er gekozen worden voor rapportages van de totaalscore en van rapportages op item-niveau, voor inzicht in foutenpatronen. Tenslotte kan met het alternatief leerlingrapport een verslag worden gemaakt voor leerlingen met een individuele leerlijn of een afwijkend toetsverloop. Leerkrachten is gevraagd van 13 verschillende rapportagemogelijkheden aan te geven welke zij zelf gebruiken (tabel 3.10). Voor de herkenbaarheid was elke rapportagemogelijkheid in de vragenlijst geïllustreerd met een afbeelding.

Tabel 3.10 Gebruik rapportagemogelijkheden

“Stel, u heeft een toets afgenomen en de toetsresultaten ingevoerd of in laten voeren. Er kunnen nu verschillende soorten rapporten gemaakt worden. Kunt u aangeven welke rapporten u gebruikt?”

	2010		2012		Verandering		
	N	Ja	N	Ja	N	-	+
Leerlingrapport							
Controle	87	53 (61%)	82	58 (71%)	67	13 (19%)	12 (18%)
Streef	74	52 (70%)	79	64 (81%)	56	6 (11%)	12 (21%)
Alternatief leerlingrapport^{ab}							
Controle ^c	87	7 (8%)	82	18 (22%)	67	5 (8%)	15 (22%)
Streef	74	14 (20%)	79	29 (37%)	56	6 (11%)	12 (21%)
Voorspelling per leerling^b							
Controle	87	5 (6%)	82	7 (9%)	67	4 (2%)	6 (3%)
Streef ^c	74	7 (10%)	79	22 (28%)	56	1 (1%)	11 (20%)
Leerlingprofiel							
Controle	87	36 (41%)	82	39 (48%)	67	10 (15%)	14 (8%)
Streef ^c	74	33 (20%)	79	48 (61%)	56	7 (13%)	18 (32%)
Categorieënanalyse leerling							
Controle ^c	87	10 (12%)	82	30 (37%)	67	2 (3%)	18 (27%)
Streef ^c	74	11 (15%)	79	38 (48%)	56	3 (5%)	16 (29%)
Groepsoverzicht – afnamemoment							
Controle ^c	86	59 (69%)	82	72 (88%)	66	5 (8%)	18 (27%)
Streef	74	50 (68%)	79	66 (84%)	56	8 (14%)	14 (25%)
Groepsrapport – gemiddelde^b							
Controle	86	8 (9%)	82	14 (17%)	66	6 (9%)	10 (15%)
Streef ^c	74	8 (11%)	79	26 (33%)	56	2 (4%)	18 (32%)
Groepsoverzicht – toets^b							
Controle	86	56 (65%)	82	56 (68%)	66	11 (17%)	11 (17%)
Streef	74	49 (66%)	79	70 (89%)	56	7 (13%)	15 (27%)
Groepsrapport – grafisch^b							
Controle	86	19 (22%)	82	23 (28%)	66	7 (11%)	14 (21%)
Streef ^c	74	19 (26%)	79	34 (43%)	56	4 (7%)	17 (30%)
Alternatief groepsprofiel^a							
Controle	86	4 (5%)	82	14 (17%)	66	4 (6%)	11 (17%)
Streef	74	10 (14%)	79	22 (28%)	56	7 (13%)	15 (27%)
Categorieënoverzicht^{bd}							
Controle	86	13 (15%)	82	20 (24%)	66	5 (8%)	7 (11%)
Streef ^c	74	8 (11%)	79	39 (49%)	56	0 (0%)	17 (30%)
Antwoordenoverzicht							
Controle	86	3 (4%)	82	4 (5%)	66	2 (3%)	3 (5%)
Streef	74	0 (0%)	79	5 (6%)	56	0 (0%)	2 (4%)
Categorieënanalyse groep							
Controle ^c	86	8 (9%)	82	24 (29%)	66	2 (3%)	13 (20%)
Streef ^c	74	5 (7%)	79	26 (33%)	56	1 (2%)	13 (23%)

= gebruikte dit rapport wel in 2010, maar niet meer in 2012; += gebruikte dit rapport niet in 2010, maar wel in 2012

^a= controlegroep en Streef-groep verschillen significant in 2010 (p<0.05)

^b= controlegroep en Streef-groep verschillen significant in 2012 (p<0.05)

^c= significant meer leerkrachten veranderen in positieve/negatieve richting (p<0.05)

^d= controlegroep en Streef-groep verschillen significant in mate van verandering (p<0.05)

Veelgebruikte rapporten zijn het leerlingrapport, groepsoverzicht-afnamemoment, groepsoverzicht-toets en leerlingprofiel⁸. In 2010 maken de leerkrachten in de Streef-groep vaker gebruik van het alternatief leerlingrapport en het alternatief groepsprofiel dan de leerkrachten in de controlegroep (respectievelijk $\chi^2(1)=4.17$, $p=0.041$; $\chi^2(1)=3.91$, $p=0.048$). In 2012 zijn de verschillen tussen de controlegroep en de Streef-groep toegenomen en maken de leerkrachten in de Streef-groep vaker gebruik van de rapporten: alternatief leerlingrapport ($\chi^2(1)=4.24$, $p=0.039$), voorspelling per leerling ($\chi^2(1)=10.2$, $p=0.001$), groepsrapport gemiddelde ($\chi^2(1)=5.41$, $p=0.020$), groepsoverzicht toets ($\chi^2(1)=9.76$, $p=0.002$), groepsrapport grafisch ($\chi^2(1)=3.95$, $p=0.047$) en categorieënoverzicht ($\chi^2(1)=10.8$, $p=0.001$).

Alle rapporten worden in 2012 vaker gebruikt dan in 2010. De controlegroep gaat in de loop der tijd vaker gebruik maken van 4 van de 13 rapporten, namelijk: alternatief leerlingrapport ($p=0.041$), categorieënanalyse leerling ($p=0.000$), groepsoverzicht-afnamemoment ($p=0.011$) en categorieën-analyse groep ($p=0.007$). In de Streef-groep worden 7 van de 13 rapporten vaker gebruikt in 2012, namelijk: voorspelling per leerling ($p=0.006$), leerlingprofiel ($p=0.043$), categorieënanalyse leerling ($p=0.004$), groepsrapport-gemiddelde ($p=0.000$), groepsrapport-grafisch ($p=0.007$), categorieën overzicht ($p=0.000$) en categorieënanalyse groep ($p=0.002$).

De rapportages die leerkrachten gebruiken, kunnen worden samengevat in de index Rapport (tabel 3.11). Hiermee wordt aangegeven in welke mate de leerkracht structureel kiest voor uitgebreide rapportage van toetsgegevens.

Tabel 3.11 Index Rapport

	2010					2012 ^b					Verandering		
	N	1	2	3	4	N	1	2	3	4	N	-	+
Controle ^c	77	3	14	46	14	79	0	14	25	40	58	10	26
		4%	18%	60%	18%		0%	18%	32%	51%		17%	45%
Streef ^c	67	1	7	34	25	75	0	5	15	55	48	6	18
		2%	10%	51%	37%		0%	7%	20%	73%		13%	38%

1= Willekeurig, niet structureel, 2= Leerling- of groepsrapport, 3= Leerling- en groepsrapport, 4= Extra foutenanalyse of alternatief profiel; -= minder uitgebreid gebruik += meer uitgebreid gebruik

^b= controlegroep en Streef-groep verschillen significant in 2012 ($p<0.05$)

^c= significant meer leerkrachten veranderen in positieve/negatieve richting ($p<0.05$)

Zowel de leerkrachten in de controlegroep als de leerkrachten die deel hebben genomen aan een Streef-project zijn in 2012 ten opzichte van 2010 meer gebruik gaan maken van de rapportagemogelijkheden van het (Cito) leerlingvolgsysteem (Controle: $p=0.011$; Streef:

⁸ De groep leerkrachten die de vragenlijst beide keren heeft ingevuld maakt vaker gebruik van het rapport 'leerlingprofiel' dan de groep die alleen in 2010 heeft geantwoord (43% vs 32%; $\chi^2(1)=4.11$, $p=0.045$); zie ook de bijlage.

$p=0.023$). In 2012 maakt bijna drie kwart van alle leerkrachten in de Streef-groep gebruik van uitgebreide rapportage. De Streef-leerkrachten maken hiermee in 2012 uitgebreider gebruik van rapportagemogelijkheden dan de leerkrachten in de controlegroep ($\chi^2(3)=9.03$, $p=0.011$).

De beschikbaarheid van toetsrapportages alleen is niet genoeg, leerkrachten moeten de uitslagen vervolgens analyseren en interpreteren. Leerkrachten is gevraagd hoe ‘breed’ het kader is waarbinnen zij naar de resultaten kijken. De index Breed kader (tabel 3.12) is samengesteld uit de antwoorden op 6 vragen naar de nadere inspectie die de leerkracht doet of de vergelijkingen die hij of zij maakt als de rapportage van de leerlingenprestaties beschikbaar is. Deze index heeft een 4-puntsschaal met oplopende complexiteit van vergelijking (geen, basis, basis plus, uitgebreid). Een basisvergelijking betekent dat leerkrachten naar eerdere toetsgegevens van leerlingen en naar de prestaties van de zwakke leerlingen kijken. ‘Basis plus’ houdt in dat leerkrachten daarnaast nog een extra analyse maken, bijvoorbeeld door naar de prestaties van de excellente leerlingen te kijken of door een vergelijking te maken met het voorgaande jaar. De classificatie ‘uitgebreid’ houdt in dat leerkrachten de toetsprestaties tegen verschillende soorten data afzetten en verschillende vakken vergelijken en een vergelijking maken met vorig leerjaar en landelijke prestaties.

In 2010 plaatsten de leerkrachten de prestatiegegevens van hun leerlingen nog niet vaak uitgebreid in een breed kader. Nog geen 5% van de leerkrachten in de controlegroep en nog geen 10% van de leerkrachten in de Streef-groep maakte een uitgebreide analyse. De leerkrachten in de controlegroep maken van 2010 tot 2012 een positieve ontwikkeling door: ruim 40% van de leerkrachten in deze groep is de toetsresultaten in 2012 in een bredere context gaan plaatsen dan in 2010 ($p=0.004$).

Tabel 3.12 Index Breed kader

	2010					2012					Verandering		
	N	1	2	3	4	N	1	2	3	4	N	-	+
Controle ^c	79	20	21	35	3	67	7	14	35	11	49	5	20
		25%	27%	44%	4%		10%	21%	52%	16%		10%	41%
Streef	64	18	13	27	6	70	19	11	29	11	43	12	12
		28%	20%	42%	9%		27%	16%	41%	16%		28%	28%

1= geen analyse of vergelijking, 2= basis, 3= basis plus, 4= uitgebreid

^c= significant meer leerkrachten veranderen in positieve/negatieve richting ($p<0.05$)

Behalve rapportages af te zetten tegen andere gegevens, kunnen leerkrachten bij de interpretatie van rapportages reflecteren op de periode voorafgaand aan de toets. De schaal Reflectie⁹ (tabel 3.13) is samengesteld uit 6 vragen over de manier waarop een leerkracht terugkijkt naar het werk en gedrag van de leerlingen, de lesstof en de toets en de eigen

⁹ Betrouwbaarheid: voormeting $\alpha=0,72$, nameting $\alpha=0,70$

werkwijze. De schaal geeft weer in hoeverre de leerkracht reflecteert, op een 3-puntsschaal van 1 (nooit) tot 3 (meestal). Zowel in 2010 als in 2012 geven beide groepen leerkrachten aan in redelijke mate te reflecteren, met een gemiddeld antwoord tussen ‘soms’ en ‘meestal’ in. De leerkrachten in de Streef-groep zijn in 2012 meer gaan reflecteren dan in 2010 ($t(48) = -2.06$, $p=0.044$).

Tabel 3.13 Schaal Reflectie

“Wilt u aangeven of u op basis van de toetsresultaten wel eens terugkijkt?”

	2010			2012			Verandering		
	N	Gem*	Sd	N	Gem	Sd	N	Gem	Sd
Controle	75	2,50	0,37	73	2,59	0,33	52	+0,05 (2,5%)	0,31
Streef ^c	68	2,45	0,35	73	2,62	0,29	49	+0,10 (5%)	0,33

*1= nooit, 2= soms, 3= meestal

^c= significante verandering van 2010 naar 2012 ($p<0.05$)

In het kader van transparantie is het belangrijk dat leerlingen zelf inzicht krijgen in wat zij (moeten) leren, wat de leerkracht van hen verwacht en hoe zij op het moment presteren. De attitude over het delen van informatie met leerlingen is gemeten met de schaal Prestatie-informatie¹⁰ (tabel 3.14), bestaande uit 3 vragen en een antwoordschaal van 1 (helemaal oneens) tot 5 (helemaal eens). Zowel in 2010 als in 2012 zijn beide groepen leerkrachten het behoorlijk eens met de stellingen over het belang van het delen van prestatie-informatie met de leerlingen, al stemde de Streef-groep er in 2010 meer mee in dan de controlegroep ($t(326) = -2.03$, $p=0.044$). Beide groepen zijn in de periode 2010-2012 meer belang gaan hechten aan het delen van prestatie-informatie met leerlingen (Controle: $t(153) = -2.66$, $p=0.009$; Streef: $t(148) = -2.84$, $p=0.005$).

Tabel 3.14 Schaal Prestatie-informatie

Stellingen over het delen van prestatie-informatie met leerlingen

	2010 ^a			2012			Verandering		
	N	Gem*	Sd	N	Gem	Sd	N	Gem	Sd
Controle ^c	171	3,92	0,64	162	4,07	0,60	154	+0,14 (3,5%)	0,64
Streef ^c	157	4,05	0,53	154	4,18	0,50	149	+0,13 (3,3%)	0,56

*1= helemaal oneens, 2= oneens, 3= niet oneens/eens, 4= eens, 5= helemaal eens

^a= controlegroep en Streef-groep verschillen significant in 2010 ($p<0.05$)

^c= significante verandering van 2010 naar 2012 ($p<0.05$)

¹⁰ Betrouwbaarheid: voormeting: $\alpha=0,71$, nameting $\alpha=0,68$

3.2 Differentiatie

Differentiatie, dat wil zeggen het omgaan met verschillende prestatieniveaus in de klas, is een belangrijke vaardigheid voor leerkrachten. Om een beeld te krijgen van de mate waarin leerkrachten differentiëren is hen gevraagd hoeveel niveaus zij voor de verschillende basisvakken hanteren en hoeveel leerlingen een individuele leerlijn hebben. Hierbij gaat het erom een haalbaar aantal niveaus te gebruiken. De antwoorden zijn samengevat in de index Niveaus (tabel 3.15). Vervolgens is gevraagd naar wat leerkrachten leerlingen laten doen als blijkt dat zij de stof niet (3 vragen) of juist heel goed hebben begrepen (4 vragen). Deze 7 vragen zijn samengevat in de index Actie. Tenslotte is in de index Attitude met 4 vragen ingegaan op de mate waarin leerkrachten positief tegenover differentiatie staan. Alle drie de indices zijn vertaald naar een 3-puntsschaal, oplopend van 1 (nooit) tot 3 (meestal), waarmee wordt aangegeven in hoeverre de leerkracht optimaal gebruik maakt van verschillende niveaus, differentieert in handelen en een positieve houding heeft ten opzichte van differentiatie.

De leerkrachten in beide groepen scoren laag tot gemiddeld op de drie indices van differentiatie. De leerkrachten in de Streef-groep lijken van 2010 tot 2012 te groeien in niveaugebruik, differentiatiepraktijken en differentiatie-attitude, maar deze veranderingen zijn niet significant. Ook de verschillen in aanpassingen van 2010 tot 2012 tussen de controlegroep en de Streef-groep zijn niet significant.

Tabel 3.15 Indices differentiatie

	2010			2012			Verandering		
	N	Gem*	Sd	N	Gem	Sd	N	Gem	Sd
Niveaus									
Controle	163	1,69	0,72	173	1,72	0,73	156	+0,05 (2,5%)	0,96
Streef	145	1,63	0,72	163	1,80	0,75	145	+0,17 (8,5%)	1,10
Actie									
Controle	80	1,98	0,66	165	1,93	0,74	75	-0,04 (2%)	0,89
Streef	67	1,87	0,69	142	1,92	0,67	55	+0,11 (5,5%)	0,76
Attitude									
Controle	171	1,94	0,74	162	1,96	0,73	154	+0,02 (1%)	0,81
Streef	155	1,88	0,80	154	2,02	0,80	149	+0,13 (6,5%)	1,00

*1= laag, 2= gemiddeld, 3= hoog

3.3 Professionaliseren en samenwerken in het team

Schoolverbeteringstrajecten hebben pas goede kans van slagen als de leerkrachten samenwerken, als team fungeren en openstaan voor verbetering van hun lespraktijk en hun ontwikkeling als leerkracht. Door middel van 2 schalen is in kaart gebracht hoe vaak leerkrachten samenwerken met collega's (schaal Samenwerking¹¹, 8 vragen, tabel 3.16) en in hoeverre zij openstaan voor hun eigen professionele ontwikkeling (schaal Beroepsattitude¹², 5 vragen). Beide schalen staan op een 5-puntsschaal, waarbij mate van Samenwerking wordt weergegeven van 1 ((vrijwel) nooit) tot 5 (1x per week) en Beroepsattitude van 1 (helemaal oneens) tot 5 (helemaal eens). Om de interpretatie van de schaal Samenwerking te vergemakkelijken, zijn de antwoorden op de afzonderlijke vragen toegevoegd (tabel 3.17).

Leerkrachten geven aan verschillende manieren van samenwerking 1 á 2 keer per jaar toe te passen (tabel 3.16). Dit is zowel het geval in 2010 als in 2012 en is vergelijkbaar voor de controlegroep en de Streef-groep. Als in meer detail naar vormen van samenwerking wordt gekeken (tabel 3.17) valt op dat regelmatige samenwerking van leerkrachten vooral plaatsvindt rond het lesmateriaal. Twee derde van de leerkrachten wisselt maandelijks tot wekelijks lesmateriaal uit met collega's en 80% praat minstens 2 keer per jaar over de ontwikkeling van nieuw lesmateriaal. Lesobservaties worden meestal niet door de leerkrachten zelf uitgevoerd: twee derde van de leerkrachten geeft aan zelf (vrijwel) nooit een collega te observeren. Conform het algemene beeld, zijn van 2010 tot 2012 weinig veranderingen te zien in mate van samenwerking. Het enige verschil is dat leerkrachten in de Streef-groep in 2012 regelmatig geobserveerd worden tijdens hun les dan in 2010 ($p=0.005$).

Tabel 3.16 Schaal Samenwerking

	2010			2012			Verandering		
	N	Gem*	Sd	N	Gem	Sd	N	Gem	Sd
Controle	85	2,47	0,58	89	2,40	0,56	68	-0,10 (2,5%)	0,56
Streef	84	2,57	0,59	77	2,54	0,64	59	+0,02 (0,5%)	0,57

*1= (vrijwel) nooit, 2= ongeveer 1x per jaar, 3= ongeveer 1x per half jaar, 4= ongeveer 1x per maand, 5= ongeveer 1x per week

¹¹ Betrouwbaarheid: voormeting $\alpha=0,69$, nameting $\alpha=0,65$

¹² Betrouwbaarheid: voormeting $\alpha=0,68$, nameting $\alpha=0,69$

Tabel 3.17 Vragen samenwerking
“Hoe vaak heeft u dit schooljaar...”

	2010						2012						Verandering		
	N	1	2	3	4	5	N	1	2	3	4	5	N	-	+
Iemand gevraagd om u te helpen bij het lesgeven van uw klas?															
Controle	82	47	11	17	3	4	89	60	8	13	5	3	66	19	9
		57%	13%	21%	4%	5%		67%	9%	15%	6%	3%		29%	14%
Streef	84	57	10	8	7	2	76	45	9	14	6	2	58	8	14
		68%	12%	10%	8%	2%		59%	12%	18%	8%	3%		14%	24%
Collega's in uw klas gehad om te observeren?															
Controle	84	26	32	23	3	0	89	30	22	31	5	1	67	17	19
		31%	38%	27%	4%	0%		34%	25%	35%	6%	1%		25%	28%
Streef ^c	84	24	28	27	5	0	77	16	17	33	11	0	59	8	25
		29%	33%	32%	6%	0%		21%	22%	43%	14%	0%		14%	42%
Bij een collega in de klas het les geven geobserveerd?															
Controle	85	58	17	6	4	0	87	58	18	10	1	0	67	9	16
		68%	20%	7%	5%	0%		67%	21%	12%	1%	0%		13%	24%
Streef	84	52	17	12	3	0	77	50	13	12	2	0	59	12	12
		62%	20%	14%	4%	0%		65%	17%	16%	3%	0%		20%	20%
Feedback op uw functioneren ontvangen van uw collega's? ^a															
Controle	84	22	26	26	6	4	88	24	24	27	10	3	67	21	15
		16%	31%	31%	7%	5%		27%	27%	31%	11%	3%		31%	22%
Streef	84	20	18	26	20	0	77	14	23	24	14	2	59	16	20
		24%	21%	31%	24%	0%		18%	30%	31%	18%	3%		27%	34%
Lesmateriaal uitgewisseld met collega's?															
Controle	84	5	1	23	38	17	87	5	4	26	33	19	65	21	14
		6%	1%	27%	45%	20%		6%	5%	30%	38%	22%		32%	22%
Streef	80	3	5	16	35	21	77	10	5	15	32	15	58	20	11
		4%	6%	20%	44%	26%		13%	7%	20%	42%	20%		35%	19%
Met collega's gepraat over ontwikkeling van nieuw lesmateriaal?															
Controle	82	3	7	25	40	7	88	10	9	27	32	10	64	25	13
		4%	9%	31%	49%	9%		11%	10%	31%	36%	11%		39%	20%
Streef	83	4	6	25	35	13	77	8	8	28	28	5	58	23	14
		5%	7%	30%	42%	16%		10%	10%	36%	36%	7%		40%	24%
Met collega's van een groep lager gepraat over wat leerlingen zouden moeten kunnen als ze bij u in de klas komen?															
Controle	57	13	13	23	7	1	65	13	23	24	3	2	43	16	11
		23%	23%	40%	12%	2%		20%	35%	37%	5%	3%		37%	26%
Streef	69	12	14	26	13	4	66	10	11	34	9	2	49	13	13
		17%	20%	38%	19%	6%		15%	17%	52%	14%	3%		27%	27%
Met collega's van een groep hoger gepraat over wat leerlingen in uw leerjaar moeten leren om in het volgende leerjaar mee te kunnen komen?															
Controle	75	18	24	20	11	2	74	23	23	19	8	1	53	20	16
		24%	32%	27%	15%	3%		31%	31%	26%	11%	1%		38%	30%
Streef	66	14	15	23	12	2	56	12	17	16	10	1	41	11	11
		21%	23%	35%	18%	3%		21%	30%	29%	18%	2%		27%	27%

1= (vrijwel) nooit, 2= ongeveer 1x per jaar, 3= ongeveer 1x per half jaar, 4= ongeveer 1x per maand, 5= ongeveer 1x per week, -= minder vaak += vaker

^a= controlegroep en Streef-groep verschillen significant in 2010 ($p < 0.05$), ^c= significant meer leerkrachten veranderen in positieve/negatieve richting ($p < 0.05$)

Voor een goede samenwerking binnen het team en een open, verbeteringsgerichte atmosfeer op school, is het belangrijk dat leerkrachten open staan voor kwaliteitsverbetering. Leerkrachten geven aan een positieve beroepsattitude te hebben en veel te doen aan hun doorlopende ontwikkeling (tabel 3.18). In 2010 waren beide groepen hier al behoorlijk positief over. De attitude van de leerkrachten in Streef-groep wordt nog positiever in 2012 ($t(56)=-4.87$, $p=0.000$) en de mate van verandering is bij deze groep is groter dan die bij de controlegroep ($t(123)=-2.57$, $p=0.011$).

Tabel 3.18 Schaal Beroepsattitude

	2010			2012			Verandering ^d		
	N	Gem	Sd	N	Gem	Sd	N	Gem	Sd
Controle	85	4,26	0,37	89	4,33	0,40	68	+0,05 (1,3%)	0,44
Streef ^c	84	4,22	0,37	75	4,45	0,40	57	+0,24 (6%)	0,36

1= helemaal oneens, 2= oneens, 3= niet oneens/eens, 4= eens, 5= helemaal eens

^c= significante verandering van 2010 naar 2012 ($p<0.05$)

^d= controlegroep en Streef-groep verschillen significant in mate van verandering ($p<0.05$)

3.4 Attitude en handelen van schoolleiders en intern begeleiders

In 2012 zijn specifiek enkele vragen gesteld aan schoolleiders en intern begeleiders, omdat een opbrengstgerichte instelling van hen een belangrijke bijdrage kan leveren aan de attitude en het handelen van de leerkrachten en het team. De gerapporteerde gegevens zijn zoals eerder vermeld de antwoorden van 120 respondenten (60 schoolleiders en 60 IB-ers), afkomstig van 75 scholen. Zij hebben gemiddeld bijna 10 jaar ervaring als schoolleider of IB-er, variërend van 1 tot 35 jaar ($M=9,8$, $sd=7,6$). Er is aan de schoolleiders en intern begeleiders (SLs/IB-ers) een selectie van de vragen gesteld, soms in een andere formulering of met andere antwoordmogelijkheden. Daardoor kunnen de schoolleiders en intern begeleiders niet vergeleken worden met leerkrachten op *alle* besproken punten uit paragraaf 3.1, 3.2 en 3.3. Desalniettemin geeft de selectie een goed beeld van de opbrengstgerichtheid van de directeuren en intern begeleiders, voornamelijk door de nadruk op hun tevredenheid over het niveau van de leerlingen in hun school, houding ten opzichte van basisvaardigheden en toetsgegevens, gebruik van toetsen en rapportages en overleg met het team. Waar dat zinvol en logisch is, worden de antwoorden van de schoolleiders en IB-ers vergeleken met de antwoorden van de leerkrachten die de vragenlijst in 2012 hebben ingevuld en actief hebben deelgenomen aan een van de deelprojecten van Streef. Dit zijn in totaal 262 leerkrachten.

3.4.1 *Tevredenheid prestaties en explicietie van doelen*

Aan de schoolleiders en IB-ers is gevraagd in hoeverre zij tevreden zijn over het niveau van de leerlingen in hun school. Deze vraag is vergelijkbaar aan een vraag aan de leerkrachten, maar is verbreed van tevredenheid op klasniveau naar tevredenheid op schoolniveau. Voor het

domein rekenen worden de antwoorden van de schoolleiders en IB-ers vergeleken met die van de leerkrachten, hoewel de vraag zoals gezegd op school- dan wel klasniveau is gericht. De gegevens over taal en lezen zijn niet goed te vergelijken, omdat deze alleen voor de leerkrachten zijn uitgesplitst naar verschillende domeinen. Daarom worden voor taal en lezen alleen de gegevens voor de groep schoolleiders en IB-ers gepresenteerd (tabel 3.19).

Tabel 3.19 Tevredenheid prestatieniveau groepen op school / leerlingen in de klas

	N	SLs/IB-ers				N	Leerkrachten			
		Ja	Nee 1	Nee 2	Nee 3		Ja	Nee 1	Nee 2	Nee 3
Rekenen ^a	116	54 (45%)	46 (40%)	10 (7%)	6 (5%)	252	91 (36%)	143 (57%)	16 (6%)	2 (1%)
Taal en lezen	116	14 (12%)	70 (60%)	25 (22%)	7 (6%)					

1= Nee, een of twee groepen (of: leerlingen) zouden beter moeten presteren; 2= Nee, zeker de helft van de groepen (of: klas) zou beter moeten presteren; 3= Nee, bijna de hele school (of: klas) zou beter moeten presteren

^a=de groepen verschillen significant van elkaar ($p < 0.05$)

De schoolleiders en IB-ers verschillen in de mate van tevredenheid over het rekenniveau ($\chi^2(3)=14,3$, $p=0.003$) van de leerkrachten. Ze zijn vaker tevreden over de hele school dan leerkrachten over hun klas zijn. Echter, 5 procent van de schoolleiders en IB-er vindt het niveau van bijna de gehele school te laag. Dit geldt zowel voor het domein rekenen als taal en lezen.

Tevredenheid over het prestatieniveau van de school hangt samen met de mate waarin er doelen zijn gesteld. Immers, als duidelijk is waarnaar gestreefd wordt, is er grond voor (on)tevredenheid over de prestaties. Aan de leerkrachten is in dit kader gevraagd in hoeverre zij het gewenste eindniveau voor hun leerlingen expliciteren. Een vergelijkbare vraag is gesteld aan de schoolleiders en IB-ers. In dit geval gaat het om de vraag in welke mate zij de doelen alleen, samen met hun collega directeur of IB-er of samen met de leerkrachten en het team expliciteren (tabel 3.20). Doordat er – anders dan bij de leerkrachten – een onderscheid is gemaakt tussen explicitering op ‘hoger’ niveau (in overleg met de directeur en de IB-er) en explicitering op teamniveau, worden de antwoorden niet vergeleken met die van de leerkrachten.

Tabel 3.20 Explicitering van doelen

	N	Geen	Zelf	Samen - hoger	Samen – team
Rekenen	103	17 (17%)	25 (24%)	16 (16%)	45 (44%)
Taal	103	17 (17%)	6 (6%)	15 (15%)	65 (63%)
Technisch lezen	100	9 (9%)	6 (6%)	12 (12%)	73 (73%)
Begrijpend lezen	104	21 (20%)	6 (6%)	14 (14%)	63 (61%)

Het vaakst worden expliciet doelen gesteld voor technisch lezen, meestal samen met het team. Opvallend is in 17 tot 20% van de directeuren en IB-ers geen doelen stelt voor rekenen, taal en begrijpend lezen. Voor het domein rekenen worden de doelen bovendien het minst vaak samen met het team geëxpliciteerd.

3.4.2 Een opbrengstgerichte attitude

De schoolleiders en IB-ers spelen een belangrijke rol in de opbrengstgerichte attitude van de school. Zoals beschreven in 3.1.1, hebben de leerkrachten vragen voorgelegd gekregen die samen een indicatie geven van de attitude over opbrengsten, op een schaal van 1 (negatieve attitude) tot 4 (positieve attitude), waarbij een neutrale houding wordt weergegeven door de waarde 2,5. Dezelfde vragen zijn voorgelegd aan de schoolleiders en IB-ers (tabel 3.21). Zowel de schoolleiders en IB-ers als de leerkrachten staan redelijk positief tegenover opbrengsten. Ze verschillen hierin niet significant van elkaar.

Opbrengstgerichtheid is vaak gericht op de basisvakken taal en rekenen. Aan de schoolleiders en IB-ers is, net als aan de leerkrachten, gevraagd hoe zij tegen de nadruk op de basisvaardigheden taal en rekenen aankijken. De schaal is samengesteld uit drie stellingen waarmee de respondent het helemaal oneens (1) tot helemaal eens (5) kon zijn. Een neutrale houding wordt weergegeven door waarde 3 (tabel 21). De leerkrachten zijn, zoals ook al gebleken was uit 3.1.1 licht positief over de nadruk op basisvaardigheden. De schoolleiders en IB-ers zijn significant positiever dan de leerkrachten ($t(261)=-3.61$, $p=0.000$), en zijn het redelijk eens met de stellingen over de nadruk op de basisvakken taal en rekenen

Tabel 3.21 Attitude opbrengsten en basisvaardigheden

	SLs/IB-ers			Leerkrachten		
	N	Gem ¹	Sd	N	Gem ²	Sd
Attitude opbrengsten ^a	116	3,04	0,85	247	2,94	0,86
Basisvaardigheden ^a	116	3,89	0,58	247	3,62	0,71

¹Op een schaal van 1 (negatieve attitude) tot 4 (positieve attitude)

²Op een schaal van 1 (helemaal oneens) tot 5 (helemaal eens)

^a de groepen verschillen significant van elkaar ($p<0.05$)

3.4.3 Bekendheid referentieniveaus

De referentieniveaus zijn zoals beschreven een middel van de overheid om een opbrengstgerichte houding in het onderwijs te stimuleren. Referentieniveaus kunnen fungeren als externe, expliciet geformuleerde doelen en kunnen daarmee bijdragen aan het stellen van expliciete prestatiedoelen in de school zelf. De referentieniveaus kunnen pas effect hebben als men er op scholen bekend mee is. Er is aan de schoolleiders en IB-ers gevraagd in welke mate zij bekend zijn met de referentie niveaus en of zij zich hier in verdiept hebben of over praten op school (tabel 3.22). Deze vraag is ook aan de leerkrachten gesteld, met het verschil dat aan de schoolleiders en IB-ers ook is gevraagd of zij over referentieniveaus praten met het schoolbestuur of met andere scholen. Om de antwoorden van de schoolleiders en IB-ers en de

leerkrachten te kunnen vergelijken is ‘overleg op schoolbestuur of met andere scholen’ samengenomen met het zich verdiepen in referentieniveaus. Praten over referentieniveaus met het team en de leerkrachten blijft geclassificeerd als de activiteit die de meeste bekendheid met referentieniveaus aangeeft, omdat dat de meest praktische invulling is van het vergroten van de bekendheid van referentieniveaus in de school.

Tabel 3.22 Bekendheid referentieniveaus

	N	1	2	3	4
SLs/IB-ers ^a	116	0 (0%)	82 (71%)	11 (10%)	23 (20%)
Leerkrachten	250	28 (11%)	113 (45%)	46 (18%)	63 (25%)

1= niet van gehoord; 2= wel eens van gehoord; 3= over gelezen of in verdiept (en over gepraat met schoolbestuur of – leiding); 4= over gesproken met team/leerkrachten.

^a=de groepen verschillen significant van elkaar ($p < 0.05$)

De schoolleiders en IB-ers verschillen significant van de leerkrachten in de mate waarin zij bekend zijn met de referentieniveaus ($\chi^2(3)=27.7$, $p=0.00$). Opvallend is dat alle schoolleiders en IB-ers weliswaar van de referentieniveaus gehoord zeggen te hebben, maar dat het daar vaak ook bij blijft. Slechts een-vijfde van de van schoolleiders en IB-ers zegt over de referentieniveaus te praten met het schoolteam.

De relatieve onbekendheid met de referentieniveaus uit zich ook in de verwachtingen van de schoolleiders en IB-ers over de invloed van de referentieniveaus op het werk van de leerkracht en de prestaties van de leerlingen (tabel 3.23). Ruim de helft verwacht geen invloed of weet het niet. De rest is overwegend positief. Slechts enkele mensen verwachten negatieve effecten. Schoolleiders en IB-ers verschillen hierin niet significant van de leerkrachten.

Tabel 3.23 Invloed referentieniveaus

	N	SLs/IB-ers			N	Leerkrachten		
		Neg	Geen	Pos		Neg	Geen	Pos
Dagelijks werk van leerkracht	116	2 (2%)	65 (56%)	49 (42%)	248	4 (2%)	123 (50%)	121 (49%)
Prestaties van leerlingen	116	0 (0%)	61 (53%)	55 (47%)	246	1 (0%)	130 (53%)	115 (47%)

Neg= negatieve invloed; geen= geen invloed of weet niet; pos=positieve invloed

Om iets meer zicht te krijgen op hoe schoolleiders en IB-ers denken over de referentieniveaus zijn 9 stellingen voorgelegd over de voor- en nadelen en het nut van referentieniveaus. De stellingen zijn ook aan de leerkrachten voorgelegd (tabel 3.24). De schoolleiders en IB-ers en de leerkrachten zijn het voor het grootste gedeelte met elkaar eens. Over drie stellingen verschillen zij van mening. Schoolleiders en IB-ers stemmen minder dan leerkrachten in met de stellingen die ingaan op de ‘oneerlijkheid’ van de referentieniveaus voor onderwijs aan zwakke leerlingen (stelling c en d; respectievelijk $t(358)=3.59$ $p=0.00$ en $t(359)=3.86$

$p=0.000$). Wellicht is dit te verklaren doordat leerkrachten zich meer direct aangesproken voelen op de prestaties van hun leerlingen en de referentieniveaus voor hen daardoor meer ‘bedreigend’ zijn. De derde stelling waarop de groepen verschillen gaat in op het gebruik van referentieniveaus bij het evalueren van toetsresultaten (stelling i). Schoolleiders en IB-ers geven aan dit minder vaak te doen dan leerkrachten ($t(360)=2.78$ $p=0.006$). Dit onverwachte resultaat is wellicht te verklaren door de selectie van de leerkrachten. Dit zijn immers leerkrachten die actief hebben deelgenomen aan een van de Streef projecten. Onderdeel van een aantal van die projecten was om leerkrachten zelf meer toetsresultaten te laten analyseren, te laten discussiëren over referentieniveaus en meetbare prestatiedoelen te stellen.

Tabel 3.24 Attitude referentieniveaus

	SLs/IB-ers			Leerkrachten		
	N	Gem*	Sd	N	Gem*	Sd
a. De meeste leerkrachten weten wel wat de leerdoelen zijn en hoe ze daar aan moeten werken.	116	3,71	0,73	244	3,66	0,71
b. Ik denk dat het vooral voor nieuwe, onervaren leerkrachten goed is als er duidelijke referentieniveaus komen.	114	3,55	0,91	245	3,58	0,96
c. Ik vind het stellen van landelijke referentieniveaus oneerlijk voor scholen met veel zwakke leerlingen ^a	116	2,53	0,79	244	2,90	0,94
d. Ik vind het stellen van landelijke referentieniveaus oneerlijk voor leerkrachten die toevallig een klas met veel zwakke leerlingen hebben ^a .	116	2,43	0,84	245	2,82	0,91
e. Ik vind het voordeel van referentieniveaus dat de leerkracht gestimuleerd wordt ook de zwakke leerlingen op een minimumniveau te brengen.	116	3,78	0,68	246	3,79	0,70
f. Ik denk dat het moeilijk wordt voor leerkrachten om te voldoen aan de referentieniveaus.	115	2,72	0,78	243	2,80	0,80
g. Ik vind het voordeel van referentieniveaus dat zij duidelijk maken wat er van leerlingen verwacht kan worden.	116	4,00	0,63	245	3,95	0,63
h. Ik denk dat referentieniveaus leerkrachten kunnen helpen om beter te differentiëren tussen leerlingen.	116	3,73	0,70	244	3,82	0,72
i. Ik maak gebruik van referentieniveaus bij het evalueren van toetsresultaten ^a .	116	2,59	0,96	246	2,89	1,00

*1= helemaal oneens; 2= oneens; 3= niet oneens/eens; 4= eens; 5= helemaal eens

^a=de groepen verschillen significant van elkaar ($p<0.05$)

Zoals ook uit de beschrijving van de opbrengstgerichtheid van schoolteams zal blijken (paragraaf 3.5), is het maken van uitgebreide rapportages van toetsresultaten niet genoeg; er moet ook naar gehandeld worden. Het gebruik maken van rapportagemogelijkheden is echter wel een begin. Er is aan de schoolleiders en IB-ers, net als aan de leerkrachten, gevraagd welke analyses zij uitvoeren (tabel 3.25). Daarnaast is aan de schoolleiders en IB-ers gevraagd

of zij gebruik maken van de module zelf-evaluatie van het Cito en zo ja, welke analyses ze gebruiken (tabel 3.26).

Tabel 3.25 Gebruik rapportage mogelijkheden

	SLs/IB-ers		Leerkrachten	
	N	Ja	N	Ja
Leerlingrapport	116	104 (90%)	124	100 (81%)
Alternatief leerlingrapport ^a	116	65 (56%)	124	39 (32%)
Voorspelling per leerling ^a	116	50 (43%)	124	37 (30%)
Leerlingprofiel	116	67 (58%)	124	67 (54%)
Categorieën analyse leerling	116	59 (51%)	124	58 (47%)
Groepsoverzicht – afnamemoment	116	98 (85%)	124	102 (82%)
Groepsrapport gemiddelde ^a	116	52 (45%)	124	39 (32%)
Groepsoverzicht – toets	116	95 (82%)	124	100 (81%)
Groepsrapport – grafisch	116	49 (42%)	124	46 (37%)
Alternatief groepsprofiel	116	37 (32%)	124	28 (23%)
Categorieënoverzicht	116	51 (44%)	124	54 (44%)
Antwoordenoverzicht	116	12 (10%)	124	8 (7%)
Categorieën analyse groep	116	39 (34%)	124	43 (35%)

^a=d e groepen verschillen significant van elkaar ($p < 0.05$)

Drie rapporten worden significant vaker gebruikt door de schoolleiders en IB-ers dan door de leerkrachten, namelijk het alternatief leerlingrapport ($\chi^2(1)=14.8$, $p=0.000$), de voorspelling per leerling ($\chi^2(1)=4.56$, $p=0.033$) en het groepsrapport gemiddelde ($\chi^2(1)=4.36$, $p=0.033$). Een groot deel van de schoolleiders en IB-ers (91 van de 115, 79%) geeft daarnaast aan gebruik te maken van de Cito zelfevaluatie. De analysemogelijkheden daarbinnen worden over het algemeen goed benut, alleen de groepsanalyse wordt wat minder vaak gebruikt.

Tabel 3.26 Gebruik mogelijkheden van Cito zelfevaluatie

	N	Ja
Dwarsdoorsnede	91	79 (87%)
Groepsanalyse	91	51 (56%)
Trendanalyse jaargroep	91	73 (80%)
Trendanalyse leerling	91	78 (86%)

Op basis van de voorgaande gegevens uit tabellen 3.25 en 3.26 komt al naar voren dat de schoolleiders en IB-ers vrij uitgebreid gebruik maken van de rapportagemogelijkheden voor toetsen. Dit blijkt ook als dergelijke informatie wordt samengevat (tabel 3.27). Driekwart van de schoolleiders en IB-ers past de een uitgebreide analyse toe. De verschillen hierin met de leerkrachten zijn niet significant.

Tabel 3.27 Index rapport

	N	1	2	3	4
SLs/IB-ers	113	1 (1%)	8 (7%)	21 (19%)	83 (74%)
Leerkrachten	117	0 (0%)	10 (9%)	27 (23%)	80 (68%)

1= Willekeurig, niet structureel, 2= Leerling- of groepsrapport, 3= Leerling- en groepsrapport, 4= Extra foutenanalyse of alternatief profiel

3.5 Opbrengstgerichtheid van teams en leerwinst bij leerlingen

3.5.1 Clusters leerkrachtenteams

Om zicht te krijgen op de vraag of en hoe opbrengstgericht werken samenhangt met leerlingprestaties, welke varianten van opbrengstgericht werken het best werken en welke elementen binnen het complexe begrip opbrengstgericht werken daaraan het meeste bijdragen wordt de samenhang tussen de opbrengstgerichtheid van teams en de leerwinst van leerlingen onderzocht. In Streef is immers het begrip opbrengstgericht werken breed opgevat, dus niet alleen het analyseren van toetsgegevens, maar ook het gebruiken en stellen van doelen, individueel en in teamverband, het vertalen van de analyse naar een gedifferentieerd groepsplan en het uitvoeren daarvan. Om te kijken welk verband er bestaat tussen opbrengstgerichtheid van schoolteams en leerwinst van de leerlingen, zijn 9 leerkracht attitudes en werkwijzen die samenhangen met opbrengstgerichtheid in kaart gebracht. De 9 kenmerken zijn:

1. mate van *samenwerking* in het team (zie 3.3, tabel 3.16)
2. mate van positieve *beroepsattitude* (zie 3.3, tabel 3.18)
3. mate van *reflectie* op leerlingprestaties en eigen werkwijze (zie 3.1.3, tabel 3.13)
4. mate waarin leerkracht *prestatie-informatie* deelt met de leerlingen (zie 3.1.3, tabel 3.14)
5. mate van positieve attitude ten opzichte van een nadruk op *basisvaardigheden* (zie 3.1.1, tabel 3.4)
6. mate van positieve attitude ten opzichte van een nadruk op *opbrengsten* (zie 3.1.1, tabel 3.3).
7. mate van *differentiatie-actie* die ondernomen wordt (zie 3.2, tabel 3.15).
8. mate waarin de meest *uitgebreide rapportagemogelijkheden* worden gebruikt (zie 3.1.3, tabel 3.11)
9. mate waarin op school wordt *gesproken over referentieniveaus* (zie 3.1.2, tabel 3.5).

Per school worden de antwoorden van alle leerkrachten samengenomen om zo een beeld van de opbrengstgerichtheid van het team te krijgen. Er is gebruik gemaakt van leerkrachtgegevens over 2010 en de prestatievoortgang van de leerlingen in de schooljaren 2010-2011 en 2011-2012. Op basis van de antwoorden van de leerkrachtenteams op de 9

kenmerken kunnen 4 patronen of clusters van schoolteams onderscheiden worden. Schoolteams die het meest **opbrengstgericht** zijn, scoren bovengemiddeld (0,5 tot 1 standaard deviatie boven het gemiddelde) op vrijwel alle meegenomen kenmerken. Deze teams werken samen, reflecteren, hechten waarde aan basisvaardigheden en opbrengsten en differentiëren ook in hun handelen. Deze leerkrachten geven dus bijvoorbeeld aan relatief vaak samenwerkingsactiviteiten te ontplooien zoals het uitwisselen van lesmateriaal met elkaar (een vraag uit de schaal ‘samenwerking’) en relatief vaak terug te kijken naar de lesstof die aan bod is geweest naar aanleiding van toetsresultaten (‘reflectie’). Zij vinden het belangrijk om veel nadruk te leggen op taal en rekenen (‘basisvaardigheden’), om altijd naar een hoger niveau te streven met de leerlingen (‘attitude opbrengsten’) en om leerlingen te vertellen wat zij de komende tijd gaan leren (‘prestatie-informatie’). Bovendien geven deze leerkrachten relatief vaak aan zichzelf te blijven ontwikkelen, bijvoorbeeld door jaarlijks te kijken hoe zij een vak beter kunnen geven en open te staan voor de werkwijze en aanpak van collega’s (‘beroepsattitude’) en ondernemen zij relatief vaak actie naar aanleiding van toetsresultaten door bijvoorbeeld zelf extra stof te zoeken voor leerlingen die de stof erg goed begrepen hebben (‘differentiatie actie’).

Daarnaast zijn er teams die wel belang hechten aan basisvaardigheden taal en rekenen en het delen van prestatie-informatie met leerlingen en die uitgebreide toets rapportages maken, maar die weinig samenwerken met collega’s, weinig reflecteren naar aanleiding van toetsresultaten en die ook weinig handelen naar de informatie uit de toets-rapportages door differentiatie toe te passen. De leerkrachten in deze teams werken **individualistisch** en zijn minder teamgericht. Een derde groep teams heeft een **behoudende** aanpak. Ze zijn niet innovatief: ze praten minder over referentieniveaus en hechten weinig belang aan extra professionalisering. Hun beroepsattitude is laag, wat betekent dat zij bijvoorbeeld aangeven minder waarde hechten aan de werkwijze en aanpak van collega’s om zelf van te leren. Ze passen echter wel differentiatie toe in hun onderwijs en geven bijvoorbeeld zowel leerlingen die goed als leerlingen die niet goed gepresteerd hebben op een toets extra of ander werk. Tenslotte zijn er schoolteams die **niet opbrengstgericht** zijn. Zij zijn gemiddeld in de mate van samenwerking en gebruik van toetsrapportages en differentiatie, maar hechten weinig belang aan basisvaardigheden en opbrengsten. Zij vinden het bijvoorbeeld minder belangrijk leerlingen te vertellen wat zij de komende tijd gaan leren, hechten minder waarde aan een nadruk op taal en rekenen en geven ook minder vaak aan altijd te streven naar een hoger niveau met hun leerlingen. De vier antwoordpatronen worden weergegeven in tabel 3.28.

Tabel 3.28 Clusters leerkrachtenteams

	Opbrengstgericht	Individualistisch	Behoudend	Niet opbrengstgericht
1. samenwerking	+	-	0	0
2. beroepsattitude	+	0	--	0
3. reflectie	+	-	0	0
4. prestatie-informatie	+	+	0	-
5. basisvaardigheden	+	+	0	-
6. attitude opbrengsten	+	0	0	-
7. differentiatie actie	+	-	+	0
8. uitgebreide rapportage	0	++	0	0
9. praten over ref.niveaus	0	0	-	0

++ is meer dan 1 sd boven het gemiddelde; + is meer dan een halve sd boven het gemiddelde; 0 is tot een halve sd boven of onder het gemiddelde; - is meer dan een halve sd onder het gemiddelde; -- is meer dan 1 sd onder het gemiddelde

De leervorderingen van leerlingen (na 1 jaar en na 2 jaar) van de *opbrengstgerichte* teams zijn vervolgens vergeleken met de drie andere typen teams. Er is per leerjaar en per inhoudsdomein gekeken of de variant van opbrengstgericht werken van het leerkrachtenteam samenhangt met leerwinst van de leerlingen. In de tabellen 3.29 en 3.30 is te zien dat de resultaten niet een volledig consistent beeld opleveren. De verwachting is immers dat de drie typen teams die afgezet zijn tegen het meest *opbrengstgerichte* team minder hoge leerwinst bij leerlingen bewerkstelligen. Deze verwachting komt grotendeels uit, hoewel niet bij elk leerstofdomein en elk cohort. Bij 61 van de 90 uitgevoerde analyses is een negatief – zij het niet altijd significant – effect zichtbaar. 12 van deze negatieve effecten zijn significant. Significante negatieve effecten geven aan dat de leerlingen van de *opbrengstgerichte* teams meer leerwinst boeken dan de leerlingen van het betreffende andere teams. Bij 29 van de 90 uitgevoerde analyses is een positief effect zichtbaar, waarvan er 3 significant zijn. Significante positieve effecten geven aan dat de leerlingen van de *opbrengstgerichte* teams minder leerwinst boeken dan de leerlingen van de betreffende andere teams. De significant negatieve effecten zijn vooral te zien bij het cluster *individualistisch*, dat gekenmerkt wordt door weinig samenwerking, reflectie en differentiatie en door een uitgebreid gebruik van analyserapportages. Dat betekent dat leerlingen bij een dergelijk team significant minder voortgang boeken dan bij een *opbrengstgericht* werkend team. Het cluster dat de karakterisering *behoudend* heeft gekregen doet nauwelijks onder voor het *opbrengstgerichte* cluster. De leerkrachten in deze teams onderscheiden zich door een lage beroepsattitude; ze zijn minder dan leerkrachten in andere teams bezig met het aanpassen van hun onderwijs, met onderwijsverbeteringen en professionalisering. Ook zijn zij minder goed op de hoogte van de referentieniveaus dan andere leerkrachten. In 3.5.2 worden per jaargroep de belangrijkste resultaten gepresenteerd.

Tabel 3.29 Effect van opbrengstgerichte teams op de leerwinst na 1 jaar^{1,2}

	Posttest in ...	Individualistisch	Behoudend	Niet opbrengstgericht
Taal kleuters	groep 2 (TvK'04)	1,40	-0,72	-3,42
Technisch lezen	groep 5	-8,34	-8,49	-6,02
Spelling	groep 3 (TvK'04)	0,17	4,59	1,75
	groep 4	1,94	0,99	-0,42
	groep 5	-0,85	-0,89	0,04
	groep 6	-0,06	-1,21	-0,49
	groep 7	0,77	1,09	0,93
Begrijpend lezen	groep 3 (TvK'04)	-7,11	-3,75	-3,28
	groep 4	1,35	-0,81	1,45
	groep 5	-2,59	-2,55	-2,27
	groep 6	-2,08	-3,99	-4,37
	groep 7	-1,93	-2,36	2,15
Rekenen	groep 2 (Ordenen)	-0,19	13,51	-0,31
	groep 3 (Ordenen)	-8,20	-3,03	-2,97
	groep 4	-2,24	-1,09	-1,43
	groep 5	-0,08	-0,42	-0,85
	groep 6	-3,52	0,25	-1,84
	groep 7	-0,14	2,26	1,26

¹ Effecten zijn geschat met meerniveau-analyses (scholen en leerlingen), en het gebruik van de term leerwinst betekent dat er voor de relevante pre-test is gecontroleerd.

² **vet**=significant effect ten gunste van opbrengstgerichte teams; *cursief*=significant effect in het nadeel van opbrengstgerichte teams

Tabel 3.30 Effect van opbrengstgerichte teams op de leerwinst na 2 jaar^{1,2}

	Posttest in ...	Individualistisch	Behoudend	Niet opbrengstgericht
Spelling	groep 3 (TvK'04)	0,13	-0,15	-0,10
	groep 4	-3,91	0,70	-0,18
	groep 5	1,07	0,41	-0,40
	groep 6	-0,93	1,05	0,88
	groep 7	-1,14	1,29	0,74
Begrijpend lezen	groep 4	-9,22	-7,74	-1,31
	groep 5	-0,36	-1,77	0,44
Rekenen	groep 3 (Ordenen)	-1,37	-0,93	3,77
	groep 4	-13,96	-4,85	-4,30
	groep 5	-6,54	-1,91	-1,07
	groep 6	1,11	0,82	-0,52
	groep 7	-3,27	-3,11	-0,65

¹ Effecten zijn geschat met meerniveau-analyses (scholen en leerlingen), en het gebruik van de term leerwinst betekent dat er voor de relevante pre-test is gecontroleerd.

² **vet**=significant effect ten gunste van opbrengstgerichte teams; *cursief*=significant effect in het nadeel van opbrengstgerichte teams

3.5.2 Resultaten per jaargroep

Groep 2

De leerlingen op scholen met *behoudende* teams boeken significant meer leerwinst op het domein rekenen dan leerlingen op scholen met *opbrengstgerichte* teams. Dit effect is een jaar later, als de leerlingen in groep 3 zitten, verdwenen. Voor taal geldt dat leerlingen van schoolteams die *niet opbrengstgericht* zijn significant minder vooruitgaan dan leerlingen van *opbrengstgerichte* schoolteams. Het is niet goed na te gaan of dit effect beklijft, aangezien in groep 3 geen vergelijkbare taaltoets wordt afgenomen.

Groep 3

De vorderingen in rekenprestaties van leerlingen van *opbrengstgerichte* schoolteams zijn groter dan de vorderingen in rekenprestaties van leerlingen van de andere typen teams. Het verschil met *individualistische* teams is hierbij significant. Dit verschil is een jaar later, als de leerlingen in groep 4 zitten, nog steeds zichtbaar. Een jaar later boeken ook de leerlingen van de *niet opbrengstgerichte* teams significant minder vooruitgang dan de leerlingen van de *opbrengstgerichte* teams op het gebied van rekenen.

Leerlingen van *behoudende* teams gaan significant meer vooruit dan leerlingen van *opbrengstgerichte* teams op gebied van spelling. Een jaar later is dit verschil verdwenen. Een jaar later blijken leerlingen van *individualistische* teams minder snel voortgang te boeken dan de leerlingen van *opbrengstgerichte* teams op gebied van spelling. Dit patroon was in schooljaar 2010-2011 nog niet significant.

Op gebied van begrijpend lezen boeken de leerlingen van *opbrengstgerichte* teams de meeste leerwinst. Het verschil met de voortgang van leerlingen van *individualistische* teams is hierbij significant. Dit verschil is een jaar later nog steeds zichtbaar. Een jaar later is ook de leerwinst leerlingen van *behoudende* teams op begrijpend leestoetsen significant kleiner.

Groep 4

In groep 4 zijn er weinig significante verschillen in voortgang in reken-, spelling- en begrijpend leesprestaties tussen leerlingen van verschillende typen teams. Alleen voor spelling is er een verschil tussen *opbrengstgerichte* teams en *individualistische* teams, in het voordeel van de laatste. Dit verschil is een jaar later, als de leerlingen in groep 5 zitten, verdwenen. Een jaar later is er wel een verschil te zien voor rekenprestaties: leerlingen van *individualistische* teams boeken minder leerwinst dan leerlingen van *opbrengstgerichte* teams. Dit patroon was al zichtbaar in schooljaar 2010-2011, maar bereikte toen nog geen significantie.

Groep 5

In groep 5 boeken de leerlingen van *opbrengstgerichte* teams vrijwel steeds meer leerwinst op de domeinen rekenen, leestempo, spelling en begrijpend lezen dan leerlingen van andere typen teams, maar deze verschillen zijn niet significant. Ook een jaar later, als de leerlingen in groep

6 zitten, zijn er geen significante verschillen tussen de leerlingen van verschillende typen teams.

Groep 6

Leerlingen van *opbrengstgerichte* teams gaan over het algemeen beter vooruit op gebied van rekenen, spelling en begrijpend lezen. De verschillen tussen vooruitgang in rekenprestaties van leerlingen van opbrengstgerichte teams en leerlingen van zowel *individualistische* als *niet opbrengstgerichte* teams zijn significant. Deze verschillen zijn een jaar later niet meer zichtbaar. Ook de verschillen in vooruitgang in begrijpend leesprestaties van leerlingen van *niet opbrengstgerichte* teams zijn significant. Er zijn onvoldoende koppelbare gegevens over vooruitgang in de begrijpend leesprestaties in groep 7 om te kunnen zien of dit effect beklijft.

Groep 7

Voor groep 7 zijn er geen significante verschillen in leerwinst tussen leerlingen van verschillende typen teams op gebied van rekenen, spelling en begrijpend lezen. Er zijn onvoldoende koppelbare gegevens over de leerlingen een jaar later, als zij in groep 8 zitten, om te zien of dit beeld zo blijft.

4 Conclusie en discussie

Streef is een verzameling schoolverbetertrajecten waarin opbrengstgericht werken en het stellen van doelen centraal staan. Met een leerkrachtvragenlijst is de attitude en het handelen rondom deze onderwijspraktijken in kaart gebracht. Er is voor gekozen om de focus te leggen op de groep van 345 leerkrachten die de vragenlijst zowel in 2010 als in 2012 heeft ingevuld. Dit is gedaan om veranderingen in de tijd te kunnen bekijken. De consequentie van deze keus is dat een deel van de data niet gebruikt is. Er zijn immers ook leerkrachten die alleen in 2010 of alleen in 2012 hebben geantwoord. Het is van belang te controleren of er geen vertekening van de resultaten optreedt door een dergelijke selectie. Dit bleek niet het geval. De geselecteerde groep leerkrachten die beide vragenlijsten heeft ingevuld verschilt nauwelijks van de leerkrachten die alleen in 2010 hebben geantwoord ($N=415$). De geselecteerde groep heeft weliswaar iets meer leservaring (gemiddeld bijna 2 jaar) en heeft iets vaker deelgenomen aan een deelproject van Streef, maar de groepen antwoorden vergelijkbaar op de verschillende vragen in de vragenlijst. Op slechts twee punten verschilden de groepen significant van elkaar: de leerkrachten die beide vragenlijsten hebben ingevuld expliciteren hun rekendoelen iets vaker samen met het team (29% vs 20%) en gebruiken iets vaker het rapport ‘leerlingprofiel’ dan de groep die alleen in 2010 heeft geantwoord (43% vs 32%). In het licht van de vele vragen die gesteld zijn, zijn deze verschillen verwaarloosbaar. Er kan dus van uit worden gegaan dat de geselecteerde groep leerkrachten die de vragenlijst zowel in 2010 als in 2012 heeft ingevuld representatief is voor de totale groep leerkrachten.

Met het huidige rapport is de stand van zaken in 2010 en 2012 weergegeven, waarbij gekeken is naar veranderingen in de tijd en naar de samenhang met deelname aan een van de Streefprojecten. Het is belangrijk om bij de interpretatie van de resultaten rekening te houden met de aard van de Streef-groep. Dit is een gemêleerde groep van 163 leerkrachten die deelnamen aan een van 4 deelprojecten die verschilden in intensiteit, inhoud en nadruk op deelonderwerpen rond opbrengstgericht werken en professionele ontwikkeling. Een effect voor de Streef-groep is daarmee een onderschatting van het effect voor de verschillende deelprojecten. Door de verschillende invulling van de deelprojecten wordt immers verandering op verschillende vragenlijstonderdelen verwacht. De resultaten op de vragenlijst voor de afzonderlijke deelprojecten worden beschreven in de deelrapportages van elk project. In de huidige rapportage wordt ingegaan op de belangrijkste algemene resultaten. Achtereenvolgens komen de volgende onderwerpen aan bod: prestatiedoelen, referentieniveaus, toetsgebruik, differentiatie en professionalisering.

Bij de vragen naar doelen en eindniveaus vallen zowel in 2010 als 2012 de problemen met begrijpend lezen op. Leerkrachten zijn het minst tevreden over het begrijpend leesniveau van

hun leerlingen (slechts 6-14% is tevreden in 2010 en 13-18% in 2012). Het tegenvallende prestatieniveau zou te maken kunnen hebben met het gebrek aan expliciete doelen voor om naar toe te werken, want leerkrachten in beide groepen hebben voor begrijpend lezen het minst vaak het gewenste eindniveau vastgesteld (13-16% in 2010; 14-27% in 2012). Op het domein rekenen zijn minder problemen. Leerkrachten in beide groepen zijn relatief tevreden met de leerlingprestaties en deze tevredenheid groeit significant van 2010 tot 2012 (25-37% in 2010 en 39-47% in 2012). Leerkrachten in de Streef-groep stellen in 2012 ook vaker rekendoelen vast in overleg met de intern begeleider of het team dan in 2010 (22% in 2010 en 42% in 2012). De Streef-interventies hanger verder samen met de opbrengstgerichte attitude van leerkrachten. Leerkrachten die deel hebben genomen aan een van de deelprojecten zijn in de periode 2012-2012 significant positiever geworden over de haalbaarheid en wenselijkheid van het verbeteren van leerlingprestaties en hebben daardoor in 2012 een significant positievere attitude over opbrengsten dan de leerkrachten in de controlegroep (score 2,9 op een schaal van 1 tot 4 waarbij 3 redelijk positief is).

Het werken met prestatiedoelen en duidelijke eindniveaus zou gefaciliteerd moeten worden door de pas ingevoerde referentieniveaus. In de periode van 2010 tot 2012 is de bekendheid met referentieniveaus bij alle leerkrachten significant toegenomen. Deze groei is sterker bij de leerkrachten in de Streef-groep en zij zijn hierdoor in 2012 bekender met de referentieniveaus dan hun collega's in de controlegroep (*onbekendheid* 21% vs 37%). Leerkrachten in beide groepen verwachten niet dan de referentieniveaus een negatieve invloed zullen hebben op hun werk en de prestaties van leerlingen, maar leerkrachten in de Streef-groep hebben wel significant vaker een positieve verwachting (68-74% vs 53%). Deze positieve verwachtingen van zijn echter getemperd in 2012. In 2012 verwachten beide groepen leerkrachten significant minder vaak een positief effect van de referentieniveaus op het dagelijks werk (Streef-groep van 74% naar 48% en controlegroep van 53% naar 38%) en verwacht de Streef-groep significant minder vaak een positief effect op leerlingprestaties van leerlingen (van 68% naar 46%). De leerkrachten zeggen in 2012 vaker geen effect te verwachten of niet te weten wat het effect zal zijn. Deze afname zou te maken kunnen hebben met de grotere bekendheid met referentieniveaus, waardoor leerkrachten een realistischer beeld hebben van de referentieniveaus en de tijd en inspanning die het kost om deze te goed integreren in de dagelijkse onderwijspraktijk. Ondanks de daling in positieve verwachting van het effect staan leerkrachten in beide groepen wel redelijk positief tegenover de referentieniveaus. Zij zien de meerwaarde vooral in het expliciteren van doelen en het gebruik bij differentiatie.

Bij het gebruik van prestatiedoelen is het afnemen en analyseren van toetsgegevens onontbeerlijk. In 2012 maken leerkrachten gebruik van meer toetsrapportages en voeren zij uitgebreidere analyses uit op de toetsgegevens dan in 2010. Dit geldt vooral voor de leerkrachten die deel hebben genomen aan een van de Streef-projecten. Zij maakten in 2010 al

significant vaker dan de leerkrachten uit de controlegroep gebruik van 2 rapportages (het alternatief leerling rapport (20% vs 8%) en het alternatief groepsprofiel (14% vs 5%)). In 2012 zijn de verschillen toegenomen en maken de Streef-leerkrachten meer gebruik van 6 van de 13 voorgelegde rapportages dan hun collega's in de controlegroep. De Streef-leerkrachten maken in 2012 dan ook significant meer gebruik van uitgebreide analysemogelijkheden dan de leerkrachten in de controlegroep (73% vs 51%). De leerkrachten in de Streef-groep zijn daarnaast in de periode 2010-2012 significant meer gaan reflecteren op hun onderwijspraktijk en de eerdere prestaties van de leerlingen naar aanleiding van toetsresultaten (5% meer, op een inschattingsschaal van 1 tot 3). De leerkrachten in de Streef-projecten hebben zich dus positief ontwikkeld in gebruik, analyse en interpretatie van toetsgegevens en zijn hiermee, in ieder geval wat betreft toetsgebruik, meer opbrengstgericht gaan werken.

Leerkrachten hebben aangegeven dat zij verwachten dat referentieniveaus kunnen bijdragen aan differentiatie. Momenteel zijn er nog verbeteringen nodig op gebied van het optimaal gebruik maken van verschillende prestatieniveaus, het aanbieden van verschillende verwerkingsactiviteiten na een toets en de houding tegenover de wenselijkheid en haalbaarheid van differentiatie. Leerkrachten in beide groepen scoren in 2010 en 2012 laag tot gemiddeld op de verschillende differentiatieonderdelen en zouden hier dus meer ondersteuning bij kunnen gebruiken.

Voor verbetering van onderwijskwaliteit is het belangrijk dat teams samenwerken, van elkaar leren en openstaan voor verandering. In de periode van 2010 tot 2012 is er volgens beide groepen leerkrachten weinig veranderd in de mate van samenwerking. De meeste samenwerking betreft het uitwisselen van lesmateriaal of het praten over de ontwikkeling van nieuw lesmateriaal met collega's. Meer dan de helft van alle leerkrachten doet dat minstens 1 keer per maand. Lesobservaties zijn een belangrijke bron van informatie bij professionele ontwikkeling. Lesobservaties worden blijkbaar meestal door de intern begeleider, schoolleider of een externe expert uitgevoerd, want maar liefst twee derde van de leerkrachten geeft aan zelf nooit een collega te observeren. Leerkrachten werken kortom in redelijke mate samen, maar er is nog ruimte voor intensivering met het oog op de professionele ontwikkeling in het team. De leerkrachten die deelnamen aan een Streef-deelproject hebben in de loop der tijd een significant positievere beroepsattitude gekregen (6% op een inschattingsschaal van 1 tot 5). Zij hebben de afgelopen periode dan ook gewerkt aan hun eigen professionalisering met behulp van Streef.

Bij onderwijsvernieuwingen zoals opbrengstgericht werken wordt vaak uitgegaan van de voortrekkende rol die schoolleiders en intern begeleiders kunnen spelen. Uit hun antwoorden op de vragenlijst blijkt dat er nog winst te behalen valt op het gebied van doelen expliciteren *samen* met het schoolteam. Voor rekenen doet ruim de helft dit (nog) niet samen met het team,

voor taal en begrijpend lezen is dit twee-derde (hoewel de leerkrachten zelf lagere percentages aangeven) en voor technisch lezen een kwart. Schoolleiders en IB-ers hebben wel een redelijk positieve opbrengstgerichte attitude, meer dan de leerkrachten. Opvallend is dat alle schoolleiders en IB-ers wel gehoord hebben van de referentieniveaus, maar dat het daar ook vaak bij blijft. Slechts een-vijfde spreekt over de referentieniveaus met het team. Dit wordt wellicht verklaard doordat de helft van de schoolleiders en IB-ers geen effecten van de referentieniveaus verwacht (of niet weet wat te verwachten). De schoolleiders en IB-ers lijken de rapportagemogelijkheden voor de Cito LOVS toetsen tenslotte goed te benutten, zowel binnen de gangbare analyses als binnen de module zelf-evaluatie van het Cito LOVS (gebruikt door 79%).

Als gekeken wordt naar de effecten van verschillende typen leerkrachtenteams kan gesteld worden dat leerlingen van *opbrengstgerichte* teams over het algemeen iets meer leerwinst boeken dan leerlingen van *niet opbrengstgerichte* en *individualistische* teams. Leerlingen van *niet opbrengstgerichte* teams gaan minder snel vooruit dan *opbrengstgerichte* teams op gebied van rekenen (groep 6), taal (groep 2) en begrijpend lezen (groep 6). Een schooljaar later zijn deze verschillen verdwenen, maar is er minder leerwinst op gebied van rekenen (groep 4 in '11-'12). Ook de voortgang van leerlingen van *individualistische* teams blijft achter in vergelijking met leerlingen van *opbrengstgerichte* teams. Zij hebben minder leerwinst op gebied van rekenen (groep 3 en 6) en begrijpend lezen (groep 3). Een schooljaar later is de leerwinst op gebied van rekenen (groep 4 en 5 in '11-'12) en begrijpend lezen (groep 4 in '11-'12) nog steeds kleiner. Leerlingen van *individualistische* teams gaan wel meer vooruit op spelling (groep 4), maar een jaar later blijft de voortgang juist achter (groep 4 in '11-'12). De leerlingen van *behoudende* teams presteren vergelijkbaar aan de leerlingen van *opbrengstgerichte* teams. Leerlingen van *behoudende* teams hebben grotere leerwinst op gebied van rekenen (groep 2) en spelling (groep 3) dan leerlingen van *opbrengstgerichte* teams. Een schooljaar later blijft de leerwinst op gebied van begrijpend lezen (groep 4 in '11-'12) van leerlingen van *behoudende* teams echter achter.

Leerlingen van *opbrengstgerichte* en *behoudende* leerkrachtenteams boeken dus meer vooruitgang. Deze teams differentiëren vaker in de klas. Een veranderingsgezinde beroepsattitude, die bij de *behoudende* teams erg laag is, lijkt niet essentieel voor het bewerkstelligen van veel leerwinst. Het kan zijn dat de leerlingen op scholen met *behoudende* teams al goede voortgang boeken en dat de leerkrachten daardoor geen noodzaak zien zichzelf te verbeteren en minder open staan voor onderwijsvernieuwingen. Daarnaast valt op dat het maken van uitgebreide toets-rapportages, zoals de *individualistische* teams doen, niet genoeg is om extra leerwinst te boeken. Leerkrachten moeten ook naar de informatie uit de rapportages *handelen* en moeten met elkaar als team werken aan en reflecteren op hun onderwijs. Dit geeft het belang aan van ondersteuning bij het schoolbreed implementeren van opbrengstgericht werken, een van de uitgangspunten van de Streefprojecten.

Samengevat kan gesteld worden dat de schoolverbetertrajecten van Streef hebben geleid tot meer en intensiever gebruik van toetsgegevens en -rapportages. Daarnaast hebben de trajecten de attitudes van leerkrachten beïnvloed, vooral op het gebied van opbrengsten en professionalisering. Beter toetsgebruik en een positievere houding tegenover aspecten van opbrengstgericht werken vormen een belangrijke basis voor kwaliteitsverbetering in het onderwijs en zijn hiermee een eerste stap in de richting van prestatieverhoging van leerlingen. Daarmee zijn scholen er echter nog niet: voor het vergroten van de leerwinst van leerlingen lijken een voortrekkende rol van de schoolleider en intern begeleider en vooral een opbrengstgericht, samenwerkend en differentiërend team essentieel.

Literatuur

- Deunk, M. I., Doolaard, S., & Hofman, R. H. (2011). *Attitude en gedrag van basisschoolleerkrachten met betrekking tot het verbeteren van leerlingresultaten. Resultaten van de beginmeting bij leerkrachten*. Groningen: GION.
- Doolaard, S. (2013). *Het streven naar kwaliteit in scholen voor primair onderwijs. Integraal eindrapport van onderzoeken in de BOPO-programmalijn onderwijskwaliteit PO 2009-2012*. Groningen: GION.
- Expertgroep Doorlopende Leerlijnen Taal en Rekenen. (2009). *Een nadere beschouwing. over de drempels met taal en rekenen*. Enschede: Expertgroep Doorlopende Leerlijnen Taal en Rekenen.
- Hattie, J. A. C. (2003). Teachers make a difference. what is the research evidence? *Australian Council for Educational Research Annual Conference on: Building Teacher Quality*. Auckland: University of Auckland.
- Hattie, J. A. C. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. London, New York: Routledge.
- Hill, H. C. (2007). Learning in the teaching workforce. *The Future of Children*, 17(1), 111-127.
- Inspectie van het Onderwijs. (2010). *Opbrengstgericht werken in het basisonderwijs*. Utrecht: Inspectie van het Onderwijs.
- Jungbluth, P. (2003). *De ongelijke basisschool: Etniciteit, sociaal milieu, sekse, verborgen differentiatie, segregatie, onderwijskansen en schooleffectiviteit*. Nijmegen / Utrecht: ITS / Sardes.
- Jungbluth, P. (2012). *Leerlingen in de ogen van hun leerkrachten*. Maastricht: Universiteit Maastricht.
- Ledoux, G., Blok, H., & Boogaard, M. (2009). *Opbrengstgericht werken. over de waarde van meetgestuurd onderwijs*. Amsterdam: SCO-Kohnstamm Instituut.
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. New York: Holt, Rinehart and Winston.
- Scheerens, J., & Bosker, R. J. (1997). *The foundations of educational effectiveness*. Oxford: Pergamon.
- Schildkamp, K., & Kuiper, W. (2010). Data-informed curriculum reform: Which data, what purposes, and promoting and hindering factors. *Teaching and Teacher Education*, 26(3), 482-496.
- Stipek, D. J., Givvin, K. B., Salmon, J. M., & MacGyvers, V. L. (2001). Teachers' beliefs and practices related to mathematics instruction. *Teaching and Teacher Education*, 17(2), 213-226.
- Visscher, A. J., & Coe, R. (2003). School performance feedback systems: Conceptualisation, analysis, and reflection. *School Effectiveness and School Improvement*, 14(3), 321-349.
- Weinstein, R. S. (2002). *Reaching higher: The power of expectations in schooling*. Cambridge, MA: Harvard University Press.

Bijlage

In deze bijlage worden de verschillen tussen de groep respondenten die de vragenlijst in 2010 en 2012 heeft ingevuld en de groep die de vragenlijst alleen in 2010 heeft ingevuld gepresenteerd.

Algemene informatie (leerkrachtgegevens in 2010):

	Alleen 2010	2010 en 2012
Aantal respondenten	415	345
Aantal verschillende scholen	90	79
Leservaring in jaren ^a (gemiddeld)	15,9	17,8
Aantal onderbouw	210 (52%)	182 (53%)
Aantal combinatiegroep	206 (50%)	162 (47%)
Deelname aan Streef ^a	278 (67%)	267 (77%)

^a De groepen verschillen significant ($p < 0.05$)

Leservaring (in 2010)

	N	Gem (in jaren)^a	Sd	Range
alleen 2010	412	15,9	11,4	0-43
2010 en 2012	345	17,8	11,1	0-40

^a De groepen verschillen significant; $t(755) = -2.26$, $p = 0.024$

Algemene deelname aan Streef

	N	Ja^a
alleen 2010	415	278 (67%)
2010 en 2012	345	267 (77%)

^a De groepen verschillen significant; $\chi^2(1) = 10.1$, $p = 0.002$

NB niet te verwarren met *actieve* deelname, zie paragraaf 2.1

Explicitering van doelen – (aanvankelijk) rekenen^a

	N	Nee	Zelf	Samen
alleen 2010	323	12 (4%)	247 (77%)	64 (20%)
2010 en 2012	274	7 (3%)	189 (69%)	78 (29%)

^a De groepen verschillen significant; $\chi^2(2) = 6.43$, $p = 0.040$

Gebruik rapportagemogelijkheden – leerlingprofiel

	N	Ja^a
alleen 2010	185	60 (32%)
2010 en 2012	161	69 (43%)

^a De groepen verschillen significant; $\chi^2(1) = 4.11$, $p = 0.045$